

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 08:09:06 ; Search time 1113 Seconds
(without alignments)

5129.900 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074

Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAGLEGLVS 1344

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPto_pool/US10066521/runat_13072004_122212_11071/app_query.fasta_1.1543
-DB=N Geneseq_29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10066521 -CGCN_1_1_819 @runat_13072004_122212_11071 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- N Geneseq_29Jan04.*
- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7074	100.0	4035	6 AAL44356	Aal44356 Human PYR
2	5756	81.4	3489	6 AAD41224	Aad41224 Human EMB
3	5756	81.4	3926	7 ADA45218	Ada45218 Human MAT
4	5741.5	81.2	3900	6 ABK48628	Abk48628 Human MAT
5	5741.5	81.2	3900	7 AAD49018	Aad49018 Human MAT
6	5724.5	80.9	5859	6 AAL47135	Aal47135 Pyrin dom
7	5724.5	80.9	6939	6 AAL47131	Aal47131 Pyrin dom
8	5724.5	80.9	6939	6 AAL47140	Aal47140 Pyrin dom

9	5683.5	80.3	3830	7 ADA45220	Ada45220 Human MAT
10	5120	72.4	3226	6 ABX97181	Abx97181 Human NOV
11	2688	38.0	3447	6 ABK48611	Abk48611 Mouse MAT
12	2688	38.0	3447	7 AAD49001	Aad49001 Mouse MAT
13	2005	28.3	1157	6 ABK48609	Abk48609 Human MAT
14	2005	28.3	1157	7 AAD48999	Aad48999 Human MAT
15	1435	20.3	3368	6 AAL44366	Aal44366 Human PYR
16	1434	20.3	2985	6 ADE36408	Ade36408 Human PAA
17	1335	18.9	1075	6 ABK48610	Abk48610 Human MAT
18	1335	18.9	1075	7 AAD49000	Aad49000 Human MAT
19	1331.5	18.8	3300	6 AAL47129	Aal47129 Pyrin dom
20	1320.5	18.7	3108	6 ADE36416	Ade36416 Human PAA
21	1316.5	18.6	3190	7 ADA53511	Ada53511 Human cod
22	1315.5	18.6	3172	6 ABN99366	Abn99366 Human sec
23	1314.5	18.6	2767	7 ABT16018	Abt16018 NOVX rela
24	1303.5	18.4	3186	6 AAL44363	Aal44363 Human PYR
25	1289.5	18.2	2575	6 AAI70683	Aai70683 Human nuc
26	1255	17.7	3857	4 AAD14323	Aad14323 Human PYR
27	1255	17.7	3857	8 ABX93556	Abx93556 Huma cDNA
28	1255	17.7	3857	8 ACD27909	Acd27909 Human PYR
29	1250.5	17.7	3531	6 AAL47128	Aal47128 Pyrin dom
30	1248.5	17.6	3189	5 AAI67185	Aai67185 Nucleotid
31	1238	17.5	2763	6 ABQ75801	Abq75801 Human MDD
32	1234	17.4	3431	4 AAD14322	Aad14322 Human nuc
33	1234	17.4	3431	8 ABX93555	Abx93555 Huma cDNA
34	1234	17.4	3431	8 ACD27908	Acd27908 Human PYR
35	1219.5	17.2	2691	6 AAL44365	Aal44365 Human PYR
36	1217	17.2	3150	5 AAI67184	Aai67184 Nucleotid
37	1189.5	16.8	3306	9 ADC30316	Adc30316 Human nov
38	1143.5	16.2	3069	6 AAL47134	Aal47134 Pyrin dom
39	1128.5	16.0	3263	6 AAI70684	Aai70684 Human nuc
40	1128.5	16.0	3459	6 ABQ78049	Abq78049 Human CGD
41	1068.5	15.1	3085	5 AAS68111	Aas68111 DNA encod
42	1001	14.2	2868	5 AAS13642	Aas13642 cDNA sequ
43	997.5	14.1	2847	4 AAS01487	Aas01487 Human sec
44	997.5	14.1	2847	7 AAD73494	Aad73494 Secreted
45	997.5	14.1	2847	7 ADA98038	Ada98038 Human sec

ALIGNMENTS

RESULT 1
AAL44356
ID AAL44356 standard; cDNA; 4035 BP.

XX AC AAL44356;

XX DT 31-OCT-2002 (first entry)

XX DE Human PYRIN-5 cDNA sequence.

XX KW Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-5.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1..4035

FT FT /*tag= a

FT FT /product= "Human PYRIN-5"

XX WO200261049-A2.

XX PD 08-AUG-2002.

XX PP 31-JAN-2002; 2002WO-US002967.

XX PR 31-JAN-2001; 2001US-0265231P.

PR 10-SEP-2001; 2001US-0318645P.
XX (MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX
XX Bertin J, Wang W, Blatcher M;
XX P-PSDB; AAO15585.
XX WPI; 2002-627477/67.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
PT diagnosing stress-related, apoptotic and inflammatory responses, or for
PT treating inflammatory and immune system disorders, cancers, or
PT neurological diseases.
XX
XX Claim 4; Fig 3; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
CC useful for modulating and diagnosing stress-related, apoptotic and
CC inflammatory responses. The PYRIN protein and DNA sequences are useful
CC for treating: inflammatory disorders and immune system disorders (e.g.
CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
CC (e.g. systemic lupus erythematosus and arthritis); and neurological
CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
CC protein and DNA sequences may also be used in screening assays, detection
CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
CC predictive medicine (e.g. diagnostic assays, clinical trials and
CC pharmacogenomics) and transcription profiling. The present DNA sequence
CC encodes the human PYRIN-5 protein
XX

SQ Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4035
Score: 7074.00 Matches: 1344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAL44356 (1-4035)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCATTTGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAAGGAAGAAATTTACAGCAATTCAGGAATTTACTAAGAAGAAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerLileProGlnPheGluLileGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCACAGTTTGAATCGAGATGCCAAGTGGAAATGCTTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerLileSerLilePheGlu 80
Db 181 CTCCTTGATGATGATTATGAGGACATCGCTGGCCCTGGGCTACGTCATTAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysLileSerGln 100
Db 241 AACATGAACCTGCGAACCCCTCTCGGAGAGGACACGGGATGACATGAAAAAATTTTCACA 300
Qy 101 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluLileSerGlnAla 120
Db 301 GCTATGNAACAAAGAGGTGCCACAGCAGCAGACAGACAGACAGAAATTTTCAAGACT 360
Qy 121 MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThr 140
Db 361 ATGGAAACAAAGAGGTGTCACAGCAGCAGACAGACAGAAACAGGACATGAGGAGTGACACA 420

Qy 141 TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArgSer 160
Db 421 TGGGACTACAGAGTACGCTGATGACCAATTCGCTGAGGAGGAGGATGTACGTCGTAGT 480
Qy 161 PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer 180
Db 481 TTTGAAAAACACTGCTGCTGACTGGCCGGAATGCAAAACGTTGGCTGGTCTTTGATTCA 540
Qy 181 AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyLileGlyLys 200
Db 541 GACCGGTGGGGCTTCGGGCTTCGACGCTGCGACGGTGGTTCACACGAAAGTCAGGAATTTGGAAA 600
Qy 201 SerAlaLeuAlaArgArgLileValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMet 220
Db 601 TCGGCTCTAGCCAGAAAGATCGTGTCTGTGGCGCAAGTGGACTCTACCAAGGGAATG 660
Qy 221 PheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSerVal 240
Db 661 TTTCTCTAGTCTTCTTCTCTCCCGTTAGAGAGATGACGCGGAAGAAGGAGACAGTGTG 720
Qy 241 ThrGluPheLileSerArgGluTrpProAspSerGlnAlaProValThrGluLileMetSer 260
Db 721 ACAGAGTTTATCTCCAGGGAGTGGCCAGACTTCCAGGCTCCGGTGCAGGAGATCATGTCC 780
Qy 261 ArgProGluArgLeuLeuPheLileAspGlyPheAspAspLeuGlySerValLeuAsn 280
Db 781 CGACCAAGAAAGGCTGTGTTCATCATGTACGGTTTCATGACCTGGGCTCTGCTCTCAAC 840
Qy 281 AsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuLileArg 300
Db 841 AATGACACAAAGCTCTGCAAAAGATCGGGCTGAGAAGCAGCTCCGTTTCACTCATACGC 900
Qy 301 SerLeuLeuArgLysValLeuLeuProGluSerPheLeuLileValThrValArgAspVal 320
Db 901 AGTCTGCTGAGGAAGGTCTCTGCTCCCTGAGTCTCTTCATGCTCCCGTCACCGTCAGACGTG 960
Qy 321 GlyThrGluLysLysLysSerGluValValSerProArgTyrLeuLeuValArgGlyLile 340
Db 961 GGCACAGAGAAAGTCAAGTCAGAGTCTGCTCTCCCGCTTACTCTGTAGTTAGAGGAATC 1020
Qy 341 SerGlyGluGlnArgLileHisLeuLeuLeuGluArgGlyLileGlyGluHisGlnLysThr 360
Db 1021 TCCGGGGAAACAAAGAAATCCACTTGTCTTTCAGCGCGGATGTTGTGAGCATCAGAAGACA 1080
Qy 361 GlnGlyLeuArgAlaLileMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla 380
Db 1081 CAAGGGTTCGTGCGATCATGAACAAACCGTGAAGTCTCGACCAAGTCCAGAGTGGCCGCC 1140
Qy 381 ValGlySerLeuLileCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla 400
Db 1141 GTGGGCTCTCTCATCTGCGTGGCCCTGCGAGCTGCGAGACGTGGTGGGGAGAGCGTCCGC 1200
Qy 401 ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro 420
Db 1201 CCCTTCAACCAACCAAGCTCACAGGCTTCGACCGCTTTTGTGTTCATCAGCTCACCCCT 1260
Qy 421 ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys 440
Db 1261 CGAGGGGTGGTTCGGGCTGTCTCAATCTCGAGAAAGAGTGTCTCTGAAGCGCTTCTGC 1320
Qy 441 ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet 460
Db 1321 CGTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAAGTGTGTGATGTGACGACCTCATG 1380
Qy 461 ValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnLileLeuLeuPro 480
Db 1381 GTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGAACATCTCTTCCCCA 1440
Qy 481 AspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla 500
Db 1441 GACAGCCACTGTGAGGAGTACTACACCTTCTTCCACTCTCAGTCTCCAGGACTTCTGTGCC 1500
Qy 501 AlaLeuTyrTyrValLeuGluGlyLeuGluLileGluProAlaLeuCysProLeuTyrVal 520

Db 1501 GCCTTGACTAGCTGTTAGAGGGCTCGAAATCGAGCCAGCTCTCTGCCCTCTGTACGTT 1560
Qy 521 GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrp 540
Db 1561 GAGAAAGACAAAGAGTCAATGGAGCTTAACAGCAGGCTTCCATATCCACTGCTTTGG 1620
Qy 541 MetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeu 560
Db 1621 ATGAAGCGTTCTTGTGTGGCTCGTGAGCGAAGCGTAAGGAGGCCACTGGAGGTCCTG 1680
Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCGCTTCCCTCGGGGTGAACAGAGCTTTCGCACTGGGTCTCTCTGTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTGAGCAGCTAATGACACCCAGGAGACACCTTGAGCGCTTCCACTGTCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTGTGTGCTTGGCATTAACAGCTTCCAAAGAGTGTGCTT 1860
Qy 621 ProLeuGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CCGATTAAACAGAACTTGAGCTTGATAGCATCTTCTCTCTGCTTCCAGCACTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyLysPheProArgAspGluSerAlaGluAla 660
Db 1921 TTGGGAAATTCGGGTGGATGTCAAGGGATCTTCCAGAGATGAGTCCGCTGAGGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 1981 TGTCTGTGGTCCCTCTATGATCGGGATAAGACCTCATTTGAGGAGCAGTGGAGAT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerIle 700
Db 2041 TTTCTGCTCCATGCTTGGCACCACCCACACCTGCGGAGCTGGACCTGGGCGAGCAGCATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLeuIle 720
Db 2101 CTGACAGCGGGCCATGAAGACCTGTGTGCAAGCTGAGGCAATCCACCTGCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCCTGATGTTTAGAATGCACAGATTACCCCTGTGTGCAGCACCTCTGGAGATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGCCACCCACCTGAAGGAAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuLeuGluSerLeuArg 780
Db 2281 GATGTAGGATGGCGTGTGAAGCTTAAACACCACCAAAATGTTTGTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysGlyLeuThrHisAlaCysThrLeuLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGGATTGCTGTGATTGACCCATGCTTACCTTGAGATCTCCCAATCTTACGACC 2400
Qy 801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCAGCCTGAAATCTCTGAGCCTGGCAGAAACAAGGTGACAGACCAGGAGTAATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTAGTGTGCTTTGAGAGTCTCCAGTGGCCCTCGCAGAAAGCTGATCTGGAGGAC 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGGCACGGTTGCAGAGTCTGAGCTTGGCCCTGAGCCCTGTCAGCAACCGGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCys 880

Db 2581 TTGACACACCTGTGCTTATCCAAACAGACCTGGGAAACGAAAGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGTGTAATCAGTGCACCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACACGCTGGCTGTGCTTCTTTCGCTATGCGCTATGCGGTAACTCATGGCTGACGCAC 2760
Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluValMet 940
Db 2761 CTGAGCCTTAGCATGAACCTGTGGAGAACATGGCGTGAAGCTTCTGTGCGAGGTCTG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuGluValLysCysHisLeuThrAlaAla 960
Db 2821 AGAGAACCATCTGTCTCATCTCCAGGACCTGGAGTGTGTAAGTGTCTATCTCACCGCGCG 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAATGCCCTGGGTGACGGTGGGGTGTCTGCGCTGTGCGAGGAGCTGAAGCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGlu 1020
Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTCTCTGATTCTCTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCACCTCTCTTGGCCCTTCTCTCAACCGGCATCTGACAGCTCTAAACCTGGTGCAGAA 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTGCGCCTTTCCTGTCGCCAGCTCTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeuGlu 1080
Db 3181 TTAACAGATAATTGGCTGTGGAATGCGACTACCTGTGCAAAATNAGGAAGCTGCTGGAG 3240
Qy 1081 GluValGlnLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAAGTCGAGCTACTCAAGCCCGAGTCGTAATTAAGCGTAGTTGGCATCTCTTTGATGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTCCGGCTCTCCCTGAAAGCCGGCATGGCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysArgValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTCTGTGGGGATGAACCCAGAGCAGAGAGGCTGTGCTTCTGTGCTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyGluSer 1160
Db 3421 TTCAAGAGCAGTACAGATTTTGGCAAGTCTCTGCTGGCCACCGCAAAATGGTAGTCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGAGTTGACAAACGTGGAGCAGAGCTCCCCCAACCCATGGCAGCAGCAACGAAACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTyrSerGlyValaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTGAGTGTGTGATATTCGGAGCCTGGTCTGAAACTGCTGAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyGlyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGGCTTGGATCCAAACAGTGTGATCATGACACCGAGGATATGGCTTGTCTACTAGGGAGA 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAAACAGTGTGTATGATGACACAGCGGTGTGTCTCTGT 3720

Qy 1241 HisTrpGluArgLeuGlySerArgGlyTrpCysLeuAsnSerAlaAspAspHisSerGly 1260
 Db 3721 CACTGGAGAGCGCTGGGCTCTAGGGGCTGGTGCTTAACAGTGTCTGATGACCCAGCGGT 3780
 Qy 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAspAsp 1280
 Db 3781 GTGTCCTGGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCACAGTGTGATGAC 3840
 Qy 1281 HisSerGlyValAlaTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSer 1300
 Db 3841 CACAGCGGTGGGCTGGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCACAGT 3900
 Qy 1301 AlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuVal 1320
 Db 3901 GCTGATGACACACAGCGGTGTGTCTGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTG 3960
 Qy 1321 SerAsnSerAlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlu 1340
 Db 3961 TCCACACAGTGTGATGACACAGCGGTGTGTCTGTGTCACTGGGAGCGGCTGGGCTCGAG 4020
 Qy 1341 GlyLeuValSer 1344
 Db 4021 GGGCTGGTGTCT 4032

RESULT 2

AD41224
 ID AD41224 standard; cDNA; 3489 BP.

AC AD41224;

DT 30-OCT-2002 (first entry)

DE Human EMBRY-1 cDNA.

Human; embryogenesis associated protein; AIDS; reproductive disorder; infertility; endometriosis; endometrial tumour; inflammatory disorder; autoimmune disorder; acquired immune deficiency syndrome; transgenic; ovarian tumour; contact dermatitis; placenta disorder; preclampsia; EMBRY-1; allergy; gene therapy; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..3489

FT /*tag= a
 FT /product= "EMBRY-1 protein"

FN WO200248362-A2.

PD 20-JUN-2002.

PF 14-NOV-2001; 2001WO-US043956.

PR 15-NOV-2000; 2000US-0249407P.

PA (INCY-) INCYTE GENOMICS INC.

PI Rankumar J, Arvizu C;

DR WPI; 2002-537629/57.

DR P-PSDB; AAE25053.

XX New polypeptides of human embryogenesis associated proteins for screening modulators useful for treating or preventing disorders e.g.

PT endometriosis, infertility, allergy, preclampsia.

PS Claim 58; Page 95-96; 97pp; English.

XX The invention relates to human embryogenesis associated proteins (EMBRY) and nucleic acid molecules encoding such proteins. EMBRY sequences are useful for screening modulators useful for treating or preventing disorders associated with abnormal expression of EMBRY. The disorders

CC treated include reproductive disorders such as infertility.
 CC endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory disorder such as acquired immune deficiency syndrome (AIDS); allergies, contact dermatitis; disorders of the placenta such as preclampsia, abruptio placentae etc. Sequences of the invention are also useful for analysing a proteome of a tissue or a cell type. EMBRY proteins are useful as immunogens for preparing antibodies. Polynucleotides of the invention are useful for creating knockin humanised animals or transgenic animals to model human diseases. They are also used in gene therapy. The present sequence is human EMBRY-1 cDNA

XX
 SQ Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3489
 Score: 5756.00 Matches: 1102
 Percent Similarity: 95.34% Conservative: 2
 Best Local Similarity: 95.16% Mismatches: 0
 Query Match: 81.37% Indels: 54
 DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AD41224 (1-3489)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTrpGlu 20
 Db 1 ATGGAAGAGAGCAAAATCGCTCACCTTTCCAGCTACGGCTGCAATGGTGTCTCTATGAG 60
 Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
 Db 61 CTAGACAGAGGAAGATTTTCAGACATTCAGAGGAATTACTAAAGAGGAATCTTCAGAAATCG 120
 Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
 Db 121 ACCACATGCTCTATTCCACAGTTTGAATCGAAGATGCCAAGCGTGAATGTCTGGCAGCTC 180
 Qy 61 LeuLeuHisGluTrpTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
 Db 181 CTCTTCATGAGTATTATGAGGATCGCTGGCTGGCTACGTCACATAGCATTAGCATTCTTGA 240
 Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLys----- 96
 Db 241 AACATGAACCTCGGAACCTCTCGGAGAGGCGGCGGATGACATGAAAAAATTCACAGAA 300
 Qy 96 ----- 96
 Db 301 GATCCTGAAGCAACGATGATGACCAAGGACCAAGCAAGAAAAAGTCCAGAAAAATAAA 360
 Qy 96 ----- 96
 Db 361 TATGGCATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAACAC 420
 Qy 97 ----- LysIleSerGlnAlaMetGluGlnGluGly 106
 Db 421 AAGTATGTTGGAATTCATTCTCTTTTTCAGAAAAATTTCAAGCTATGGAACAAGAGGT 480
 Qy 107 AlaThrAlaAlaGluThrGluGlnGluIleSerGlnAlaMetGluGlnGluGlyAla 126
 Db 481 GCCACAGCAGCAGACAGACAGAAACAAGAAATTTCAAGCTATGGAACAAGAGGTGCC 540
 Qy 127 ThrAlaAlaGluThrGluGlnGlnGlyHisGlyValPheThrTrpAspTrpLysSerHis 146
 Db 541 ACAGCAGCAGACAGACAGACAGACAGCATGGAGGTGACACATGGGACTACAGAGTCCAC 600
 Qy 147 ValMetThrLysPheAlaGluGluAspValArgArgSerPheGluAsnThrAlaAla 166
 Db 601 GTGATGACCAAAATTCGCTCAGGAGGAGGATGTACGTCGTAGTGTTCGAAAAACACTGCTGCT 660
 Qy 167 AspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArg 186
 Db 661 GACTGCGCGGAATGCAACGTTGGCTGTGCTTTTGATTCAGACCGGTGGGCTTCCCG 720
 Qy 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArg 206

Db 721 CCTCCACGGTGGTTCTGACGGAAAGTCAGGAATTGGGAATCGGCTCTAGCCAGAAGG 780
Qy 207 lleValLeuCystrpAlaGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 781 ATCGTGTCTGTGCTGGGCGCAAGGTGGATCTTACCAAGGAAATGTTCTCTACGTCTTCTTC 840
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 841 CTCCCGGTAGAGAGATCAGCGGAAGAAGAGAGACAGTGTACAGAGTTCATCTCCAGG 900
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 901 GAGTGGCCAGACTCCCGAGCTCCCGTGACGAGATCATGTCCCGACCAAGAAAGGCTGTG 960
Qy 267 PheIleIleAspGlyPheAspLeuGlySerValLeuAsnAsnAspThrLysLeuCys 286
Db 961 TTTCATCATGACGGTTTCGATGACTGGGCTCTGTCTCAACAATGACACAAGCTCTGC 1020
Qy 287 LysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeuArgLysVal 306
Db 1021 AAAGACTGGGCTGGAAGACAGCTCCGTTCACTCATACGACGTCTGCTGAGGAAGGTC 1080
Qy 307 LeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLys 326
Db 1081 CTGCTCCCTGAGTCTCTCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAG 1140
Qy 327 SerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGlnArgIle 346
Db 1141 TCAGAGGTCTGTCTCCCGTTACTCTGTATTAGAGAAATCTCGGGGAACAAGAAATC 1200
Qy 347 HisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIle 366
Db 1201 CACTTGCTCCTTGAGCGCGGATTTGGTGAGCATCAGAAAGACACAAAGGTTCCGTCGATC 1260
Qy 367 MetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCys 386
Db 1261 ATGAACAACCGTGAGCTCTCGACAGTGCAGAGTGGCCGCGCTCTCTCATCTGC 1320
Qy 387 ValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db 1321 GTGGCCCTGAGCTGCAGAGAGTGGTGGGAGAGCGTCCGCCCTTCACCAAAACGCTC 1380
Qy 407 ThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArg 426
Db 1381 ACAGGCTGCACGCGCTTTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGTCCGCGCG 1440
Qy 427 CysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGly 446
Db 1441 TGTCCTCAATCTGGAGGAAGAGTTGTCTGAAGCGCTTCTGCCCGTATGGCTGTGGAGGA 1500
Qy 447 ValTrpAsnArgLysSerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGlu 466
Db 1501 GTGTGGATAGGAAGTCAGTGTTGACGGTGACGACCTCATGTTCAAGGACTCGGGGAG 1560
Qy 467 SerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGlu 486
Db 1561 TCTGAGCTCCGTGCTCTGTTTCACTGAACATCTCTCTCCAGACAGCCACTGTGAGGAG 1620
Qy 487 TyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeu 506
Db 1621 TACTACACTTTCTCCACTCAGTCTCCAGAGCTTCTGTGCCGCTTGTACTAGCTGTTA 1680
Qy 507 GluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSer 526
Db 1681 GAGGCGCTGGAAATCGACCGAGCTCTCTGCCCTCTGTACGTTGAGAGACAAGAGGTCC 1740
Qy 527 MetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPhe 546
Db 1741 ATGGAGCTTAAACAGCGAGGCTTCCATATCCACTCGCTTTTGGATGAAGCGTTTCTGTGTT 1800
Qy 547 GlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValPro 566
Db 1801 GGCCTCGTGAGCGAAGACGTAAAGAGGCGCACCTGGAGGTCCTGTGGGCTGTCCCGTTCCC 1860

Qy 567 LeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAla 586
Db 1861 CTGGGGGTGAAGCAGAGACTTTCGACTGGGTCTCTCTGTGGGTGAGCAGCTTAATGCC 1920
Qy 587 ThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGlu 606
Db 1921 ACCACCCAGGAGACACCTTGGACGCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAG 1980
Qy 607 PheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeu 626
Db 1981 TTTGTTTCGCTTGGCATTAACAGCTTCCAAAGAGTGTGGCTTCGATTAACAGAACCTG 2040
Qy 627 AspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgVal 646
Db 2041 GACTTGATGACATCTTCTTCTGCTCCAGCACTGTCCGATTTTTCGGAATAATTCGGGTG 2100
Qy 647 AspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeu 666
Db 2101 GATGTCAAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCATGTCTGTGGTCCCTCTA 2160
Qy 667 TrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGly 686
Db 2161 TGGATCGGGATTAAGACCTTCATTTAGGAGCAGGTGGGAAGATTTCTGCTCCATGCTTGGC 2220
Qy 687 ThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMet 706
Db 2221 ACCCACCCACACTGCGGCAAGCTGGACCTGGGCAGCAGCATCTCTGACAGAGCGGCCATG 2280
Qy 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
Db 2281 AAGACCTGTGTGCAAGCTGAGGCATCCCACTGCAAGATACAGACCTGTATGTTTTAGA 2340
Qy 727 AsnAlaGlnThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsn 746
Db 2341 AATGCACAGATTAACCTTGTGTGACGACCTCTGGAGAAATCGTCATGGGCCAACCGTAAC 2400
Qy 747 LeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCys 766
Db 2401 CTAAGATCCCTCAACTTGGGAGGACCCACCTGAGGAAGAGGATGTAAAGATGGCGTGT 2460
Qy 767 GluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyLeu 786
Db 2461 GAAGCCTTAAACACACCCAAATGTTTGTGGAGTCTTTGAGGCTGGATGTGTGGATG 2520
Qy 787 ThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrSerProSerLeuLysSer 806
Db 2521 ACCCATGCTGTTTACCTGAAGATCTCCAAATCTTACGACCTCCCGCAGCTGAAATCT 2580
Qy 807 LeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeu 826
Db 2581 CTGAGCTTGCAGGAAACAAGGTGACAGACAGGGAGTAATGCTCTCAGTGATGCCCTTG 2640
Qy 827 ArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThr 846
Db 2641 AGAGTCTCCAGTCCGCGCTCGCAAGACTGATACTTGGAGGACTGTGGCATCACAGCCACG 2700
Qy 847 GlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeu 866
Db 2701 GGTGTCCAGAGTCTGGGCTCAGCCCTCGTCAGCAACCGGAGCTTGAACACACCTGTGCCCTA 2760
Qy 867 SerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuPro 886
Db 2761 TCCAAACAACAGCTTGGGAACGAAGGTGTAAATCTACTGTGTGATCCATGAGGCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCACCTGGACACGGCTGGCTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGCACCTTGGCTTATGGGTAATCTCATGTGCTGACGCACTGAGCCTTAGCATGAAC 2940

QY 927 ProValGluAspAenGlyValIysLeuLeuLeuCysGluValMetArgGluProSerCysHis 946
 DB 2941 CCTGTGGAGACATGGGTGAGCTTCTGTGGAGGTCATGAGAGAACCATCTGTTCAT 3000
 QY 947 LeuGlnAspLeuLeuValIysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966
 DB 3001 CTCAGGACCTGGAGTGGTAAAGTGTCTATCCACCGCGGTGCTGTGAGAGTCTGTCC 3060
 QY 967 CysValIleSerArgSerArgHisLeuIysSerLeuAspLeuThrAspAsnAlaLeuGly 986
 DB 3061 TGTGTGATCTCGAGAGCAGACACCTGAAGAGCCTGGATCTTCACGGACAATGCCCTGGT 3120
 QY 987 AspGlyGlyValAlaAlaLeuCysGluGlyLeuIysGlnIysAsnSerValLeuThrArg 1006
 DB 3121 GACGGTGGGTGTGCACTGTGGAGGACTGAAGCAAAAGAACAGTGTTCGACGAGA 3180
 QY 1007 LeuGlyLeuIysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
 DB 3181 CTCGGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGTGAGGCACCTCTCCTGGCCCTT 3240
 QY 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProIysGly 1046
 DB 3241 TCCTGCAACCGGCATCTCACAGCTAAACCTGTGTGCAGATAAATCTTCAGTCCCAAGGA 3300
 QY 1047 MetMetIysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleGlyLeu 1066
 DB 3301 ATGATGAAGCTGTGTTCGGCTTTTGCCTGTCACAGTCTAACTTACAGATAATTTGGGCTG 3360
 QY 1067 TrpIysTrpGlnTyrProValGlnIleArgIysLeuLeuGluValGlnLeuIys 1086
 DB 3361 TGGAAATGGCAGTACCTCTGCCAAATGAAGAAAGTGTGTGGAGGAAGTGCAGTACTCAAG 3420
 QY 1087 ProArgValValIleAspGlySerTrpHisSerPheAspGluAspArgHis 1104
 DB 3421 CCCGAGTCGTAATGACGGTAGTTGGCAATCTTTTGTATGAAGATGACCGGTAC 3474
 RESULT 3
 ADA45218
 ID ADA45218 standard; DNA; 3926 BP.
 XX
 AC ADA45218;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human MATER DNA.
 XX
 KW human ; MATER; maternal antigen that embryos require; NTP-ase;
 KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;
 KW antinflammatory; immunosuppressive; gynaecological; endometriosis;
 KW ovarian dysfunction; autoimmune premature ovarian failure;
 KW autoimmune disease; contraceptive; autoantibody; ds; gene.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3489
 FT /tag= a
 FT /product= "MATER"
 XX
 XX EP1285964-A2.
 XX
 XX 26-FEB-2003.
 PD
 XX
 PF 12-JUL-2002; 2002EP-00090246.
 XX
 FR 10-AUG-2001; 2001DE-01039874.
 XX
 PA (SCHD) SCHERING AG.
 XX
 FI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
 XX
 DR WPT; 2003-302814/30.
 DR P-PSDB; ADA45219.

XX New nucleic acid for human maternal antigen that embryos require protein,
 PT useful e.g. in diagnosis and treatment of female infertility.
 XX
 PS Claim 1; Page 9-11; 31pp; German.
 XX
 CC This invention describes a novel human MATER (maternal antigen that
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
 CC associated with apoptosis, a defect in MATER activity causes growth
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
 CC MATER gene is found predominantly in ovary, testis and placenta. The
 CC products of the invention have antiinfertility, contraceptive,
 CC antiinflammatory, immunosuppressive and gynaecological activity.
 CC Effectors of the MATER polypeptide are used to treat infertility
 CC associated with endometriosis and also ovarian dysfunction, autoimmune
 CC premature ovarian failure, inflammation, autoimmune diseases and female
 CC infertility, and as contraceptives. Measurements of autoantibodies
 CC against MATER in body samples or MATER protein or mRNA in eggs is useful
 CC for diagnosis of female infertility, particularly by detecting mutations
 CC in the MATER gene with a DNA chip. This sequence encodes the human MATER
 CC polypeptide described in the invention.
 XX
 SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3926
 Score: 5756.00 Matches: 1102
 Percent Similarity: 95.34% Conservative: 2
 Best Local Similarity: 95.16% Mismatches: 0
 Query Match: 81.37% Indels: 54
 DB: 7 Gaps: 1
 US-10-066-521-6 (1-1344) x ADA45218 (1-3926)
 QY 1 MetGluGlyAspIysSerLeuThrPheSerSerTrpGlyLeuGlnTrpCysLeuTrpGlu 20
 DB 1 ATGGAAGGAGACAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGGTGTCTCTATGAG 60
 QY 21 LeuAspIysGluGluPheGlnThrPheIysGluLeuLeuIysIysSerSerGluSer 40
 DB 61 CTAGACAAGGAAGAATTTTCAGACATTTCAAGGAATTTACTAAGAAAGAAATCTTCAGAAATCG 120
 QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
 DB 121 ACCACATGCTCTATTCCACAGTTTGAATTCGAGATGCCACGTCGGAATGCTCGGACATC 180
 QY 61 LeuLeuHisGluTrpTrpGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
 DB 181 CTCTTGCATGAGTATTATGAGCATCGCTGGCCTGGGCTACGTCCATTTAGCATCTTTTGA 240
 QY 81 AsnMetAsnLeuArgThrLeuSerGluIysAlaArgAspAspMetIys----- 96
 DB 241 AACATGAACCTGCAACCCCTCTCGGAGAAGGCCGCGGATGACATGAAAAAATTCACCAGAA 300
 QY 96 ----- 96
 DB 301 GATCCTGAGACACGATGACTGACCAAGGACCAAGCAAGAAAAAGTGCAGAAAAATAAA 360
 QY 96 ----- 96
 DB 361 TATGGCATGACTAAGCTTATCTTTGGGGTGTCTGACATCTCTGACTCGAATAATAAACAC 420
 QY 97 -----LysIleSerGlnAlaMetGluGlnGly 106
 DB 421 AAGTATGTTGGAATTCATTCTTTTTCAGAAAATTTTCAAGCTATGGAACAAGAGGT 480
 QY 107 AlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAlaMetGluGlnGlyAla 126
 DB 481 GCCACAGCAGCAGACAGACAGAGAAACAAAGAAATTTCAAGCTATGGAACAAGAGGTGCC 540
 QY 127 ThrAlaAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTrpIysSerHis 146
 DB 541 ACAGCAGCAGACAGACAGAAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCCAC 600


```
Db 2761 TCCACACAGCCTGGGAAACGAGGTGTAATCTACTGTGTGATCCATGAGGCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGCTGATGCTGAATCAGTGCCACCTGGACACGGCTGGCTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGACCTTGGCTTATCGGTAACTCATGGCTGACGCACCTGAGCCTTAGCATGAAC 2940
Qy 927 ProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
Db 2941 CCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTCAATGAGAAACCACTGTGTCAT 3000
Qy 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966
Db 3001 CTCAGGACCTTGGAGTTGGTAAAGTGTCTATCTACCGCGCGTGTGTGAGAGTCTGTCC 3060
Qy 967 CysValLysSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
Db 3061 TGTGTGATCTCGAGAGCAGACACTGAAGAGCCTGGATCTCACGGACAATGCCCTGGGT 3120
Qy 987 AspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArg 1006
Db 3121 GACGGTGGGTGTGCTGCACTGTGCGGAGTGAAGCAAGAACAGATGTTCTGACGAGA 3180
Qy 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
Db 3181 CTCGGTTGAAGGCATGTGGACTGACTTCTGATTTGCTGTGAGGCACCTCTCCTGGCCCT 3240
Qy 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGly 1046
Db 3241 TCCTGCAACCGGCATCTGACAGTCTAACTCTGTGCGAGATACTTCAGTCCCAAGGA 3300
Qy 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnLysLeuGlyLeu 1066
Db 3301 ATGATGAAGCTGTGTCGGCTTTGCGCTGTCACAGTCTAACTTACAGATAATTGGCGTG 3360
Qy 1067 TrpLysTrpGlnTrpProValGlnLeuArgLysLeuLeuGluValGlnLeuLys 1086
Db 3361 TGGAAATGGCATGACCTGTGCAATAGGAAGCTGCTGGAGGAAGTGCAGTACTCAAG 3420
Qy 1087 ProArgValValLysAspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3421 CCCCAGTCTGTAATGACGGTAGTGGCAATCTTTTGATGAAGATGACCCGGTAC 3474
```

RESULT 4

```
ABK48628 standard; cDNA; 3900 BP.
ID ABK48628;
AC ABK48628;
XX
XX
XX 13-AUG-2002 (first entry)
XX Human MATER cDNA.
XX
XX Human; gene; ss; contraceptive; antinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..3603
CDS /*tag= a
FT /product= "Human MATER"
FT
XX
XX WO200232955-A1.
XX
XX 25-APR-2002.
XX
```

PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2002-454595/48.
DR P-PSDB; AAU79526.
XX
XX New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
XX Claim 11; Page 80-85; 93pp; English.
XX
XX The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA.
CC Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to
CC published human genomic DNA sequences, while the remainder of the
CC sequence was determined by direct cloning of human ovarian cDNAs
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 5741.50 Matches: 1098
Percent Similarity: 96.16% Conservative: 3
Best Local Similarity: 95.90% Mismatches: 3
Query Match: 81.16% Indels: 41
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x ABK48628 (1-3900)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyGlyLeuGlnTrpCysLeuTyGlu 20
Db 154 ATGGAAGGAGACAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGGTCTCTATGAG 213
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysLysSerSerGluSer 40
Db 214 CTAGCAAGGAAGAAATTTACAGATTTCAAGGAATTTACTAAGAGAAATCTTCAGATCG 273
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 274 ACCACATGCTCTATTCCACAGATTTGAAATCGAGAATGCCAACCTGGAATGTCTGGCACC 333
Qy 61 LeuLeuHisGluTyTrpGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 334 CTCTTCATGATGAGTATTATGAGCATGCTGGCTCGGCTACGTCCTCATTTAGCATCTTTGAA 393
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
Db 394 AACATGAACCTGCAACCCCTCTCGAGAGAGCGACCGGATGACATGAAGACATTCACCA 453
Qy 97 ----- 97
Db 454 GAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAAGTGCCGGAATT 513

QY 98 -----IleSer 99
Db 514 TCACAAGCTGTGCAACAAGATAGTCCACAGCTGCAGAGACAAAAGAACAGGAAATTTC 573
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlySerGln 119
Db 574 CAAGCTATGGAAACAGAAAGTCCACAGCAGCAGAGACAGAAAGAACAAAGAAATTTCACAA 633
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAsp 139
Db 634 GCTATGGAAACAAGAGGTGTCACACAGCAGCAGACAGACAGAAAGACATGGAGGTGAC 693
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
Db 694 ACATGGGACTACAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGATGACGTGCT 753
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db 754 AGTTTGTGAAACACATGCTGCTGACTGGCCGGGAAATGCAAAACGTGGCTGGTCTTTTGAT 813
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGGCTTCGGGCTTCGACGGTGGTCTTCGACGGAAAGTCAGGAATTGGG 873
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlnGlyLysTrpGlnGly 219
Db 874 AAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGCGCGCAAGGTGGACTCTTACCAGGGA 933
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTCACTGCTCTTCTTCCTCCCGTTAGAGAGATGCGACGGGAAGAGAGACAGT 993
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTCCACAGAGTTTCATCTCCAGGGAGTGGCCAGACTCCCGAGGCTCCGGTACGGAGATCATG 1053
QY 260 SerArgProGluArgLeuLeuPheIleLeAspGlyPheAspLeuGlySerValLeu 279
Db 1054 TCCCGACCAAGAGCTGTGTGTTTCATCTGACGGTTCGATGACCTGGGCTCTGTCTCTC 1113
QY 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
Db 1114 AACATATGACACAAAGCTCTGCAAGACTGGGCTGAGAGACGCTCCGTTCCACCTCATAT 1173
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 CGCAGTCTGCTGAGGAAGTCTGTCTCCCTGATCCTTCTGATCTCACCCTCAGAGAC 1233
QY 320 ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGCACAGAGAAGCTCAAGTCAAGGTCGTGTCTCCCGGTACCTCTGTAGTTAGAGA 1293
QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuArgGlyIleGlyGluHisGlnLys 359
Db 1294 ATCTCCGGGGAAACAAGAAATCCACTTGTCTCTTGAAGCGGGATTTGGTGCATCAGAAG 1353
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
Db 1354 ACACAAGGGTTCGTGCCATCATCAACCAACCTGTAGCTGCTCGACCAAGTGCAGGTGCC 1413
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399
Db 1414 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAGAGCGTC 1473
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
Db 1474 GCCCCCTTCAACCAACCGCTCACAGGCTGCAGCGCGCTTTTGGCTTTTCATCAGCTCACC 1533
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLysArgPhe 439
Db 1534 CCTCAGGCGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGATGTCTCCTGAAGCGCTTC 1593

QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
Db 1594 TGCCTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTGATGGTGCACACCTC 1653
QY 460 MetValGlnGlyLeuGlyGlySerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1654 ATGGTTCAAAGACTCCGGGAGTCTGAGCTCCGTCTCTGTTTCACATGAATCATCTCTCTC 1713
QY 480 ProAspSerHisCysGluGluTyrTrpPhePheHisLeuSerLeuGlnAspPheCys 499
Db 1714 CCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGT 1773
QY 500 AlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
Db 1774 GCCGCTTGTACTACGTGTAGAGGGCTGGAATCGAGCCAGCTCTCTGCCCTCTGTAC 1833
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
Db 1834 GTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGACGGCTTCCATATATCCACTCGCTT 1893
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTTCTGTGTGGCTCTGTGAGCGAAGACGTNAGAGGCCACTCGAGGTC 1953
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTG 2013
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGGTGAGCAGCGCTTAATGCCACCCACAGGAGACACCTTGGACGGCTTCCACTGTCTT 2073
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
Db 2074 TTCAGACTCAAGCAAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTCCAAAGAAGTGGG 2133
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCCGATTAACACAGAACCTGGACTGTAGTACATCTTCTTCTGCTCCAGCAGCTGTCCG 2193
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
Db 2194 TATTTGGGAAAAATTCGGGTGGATGTCCAAAGGGATCTTCCCAAGAGATGATCCGCTGAG 2253
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
Db 2254 GCATGCTCTGTGTCCTCTATGGATCGGGATNAGACCTCATTTGAGGAGCAGTGGGAA 2313
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCTGCTCCATGCTTGGCACCCACACCTCGCGCAGCTGGACCTGGGCAGCAGC 2373
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCACTCCACCTGCAAG 2433
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGTGATGTTTAGAAATGCACATTAACCTCGTGTGCAACACCTCTGAGAA 2493
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 2494 ATCGTCATGCCAACCGTAACCTAAGATCCCTCAACTTGGGAGGACCCACCTGAGAGGA 2553
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGAAGGATGGCGTGTGAAGCCTTAAACACCCAAAATGTTTGTGGAGTCTTTG 2613
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGATGCTGTGATGATGACCCATGCTGTACCTGAAGATCTCCCAAAATCTTTACG 2673
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819

```
Db 2674 ACCTCCCCACCGCTGAATCTCTGAGCCTGCAGGAAACAAGGTGCAGACCCAGGAGTA 2733
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 2734 ACGCCTCTCAGTGATGCTTCCAGGCTCTCCAGTGCCTGCAGAGCTGATACTGAG 2793
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2794 GACTGTGGCATCACAGCCACGGTTCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 2853
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2854 AGCTTGACACACCTGTGGCTATCCACACAGCCCTGGGGAACGAGGTGTAAATCTACTG 2913
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2914 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCAC 2973
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2974 CTGGACACGGCTGGCTGTGCTTCTTGCACCTTGCCTTATGGGTAACTCATGTGCTGACG 3033
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 3034 CACCTGAGCCTTAGCATGAACCTTGTGGAGACAAATGGCGTGAAGCTTCTGTGGCAGGTC 3093
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 3094 ATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCAATCTCACCGCC 3153
Qy 960 AlaCysCysGluSerLeuSerCysValLysSerArgSerArgHisLeuLysSerLeuAsp 979
Db 3154 GCGTGTCTGAGAGTCTGCTCTGTGTGATCTCGAGGAGACACACCTGGAAGAGCTTGAT 3213
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 3214 CTCACGGCAATGCCCTGGGTGAGCGTGGGTGCTGCGCTGTGCGAGGAGCTGAAGCAA 3273
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 3274 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGAGCTGACTTCTGATTGCTGT 3333
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 3334 GAGGCACCTCTCCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTCTAAACCTGGTGACG 3393
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 3394 AATAACTTCAGTCCCAAGGAATGATGAGCTGTGTTGGCCTTGGCTGCCACGCTCT 3453
Qy 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 3454 AACTTACAGATAATTGGGCTGTGAAATGGCAGTACCTGTCGCAATAAGGAAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAAGTGCACTACTCAAGCCCGAGTCTGTAATTGACGGTAGTGTGGCATCTCTTTTGTAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588
```

RESULT 5

```
AAD49018
ID AAD49018 standard; cDNA; 3900 BP.
XX
AC AAD49018;
XX
DT 07-MAR-2003 (first entry)
XX
DE Human MATER cDNA.
XX
Qy 97 ----- 97
```

Human; MATER protein; infertility; fertility; contraceptive agent;
gene therapy; gene; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 1..3603
/tag= a
/product= "Human MATER protein"

WO200281492-A1.

17-OCT-2002.

29-MAR-2002; 2002WO-US009776.

04-APR-2001; 2001WO-US010981.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Nelson LM, Tong Z;

WPI; 2003-058494/05.

P-PSDB; AAE31749.

New isolated variant MATER proteins and nucleic acids, useful for
diagnosing, prognosing and treating infertility and reduced fertility,
and as contraceptive agents.

Example 2; Page 93-98; 110pp; English.

The present invention relates to novel MATER proteins and polynucleotides
encoding such proteins. The MATER proteins are essential to fertility.
Sequences of the invention are useful for diagnosing, prognosing and
treating infertility, reduced fertility and as contraceptive agents. They
are also useful in gene therapy. The method is useful for detecting a
predisposition to or pre-symptomatic screening of an individual for
infertility or reduced fertility. The present sequence is human MATER
cDNA

Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3900
Score:	5741.50	Matches:	1098
Percent Similarity:	96.16%	Conservative:	3
Best Local Similarity:	95.90%	Mismatches:	3
Query Match:	81.16%	Indels:	41
DB:	7	Gaps:	1

US-10-066-521-6 (1-1344) x AAD49018 (1-3900)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTrpGlu 20

Db 154 ATGAAGGAGACAAATCGCTCACCTTTCAGCTACGGGCTGCAATGGTCTCTATGAG 213

Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40

Db 214 CTAGACAAGGAAGAAATTTAGACATTCAGGAATTTACTTAAGAGAAATCTTCAGAAATCG 273

Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60

Db 274 ACCACATGCTCTATTTCACAGTTTGAATCGAATGCGCAATGCCAATGCGTCTGGCACTC 333

Qy 61 LeuLeuHisGluTyrTrpGlyAlaSerLeuAlaThrSerIleSerIlePheGlu 80

Db 334 CTCTTCATGAGTATTATGAGGACATCGCTGGCTGGCTACGTCCTAGTATCTTTGAA 393

Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97

Db 394 AACATGAACCTGCGAAACCTCTCTCGAGAGAGGACCGGGATGACATGAAGACATTACCA 453

Db 454 GAAGATCCTGAAGCAACGATGATGACCAAGGACCAAGCAAGGAAAAAGTCCAGGAATT 513
QY 98 -----IleSer 99
Db 514 TCACAAGCTGTGCAACAAGATAGTCCACAGCTGCAGAGACAAAGAAACACAGAAATTTC 573
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluSerGln 119
Db 574 CAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGAAATTTCACA 633
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlyHisGlyAsp 139
Db 634 GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGGACATGGAAGTGAC 693
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg 159
Db 694 ACATGGGACTACAGAGTCAGTGTATGCCAAATTCGCTGAGGAGGAGATGACGTGCT 753
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db 754 AGTTTGAACACATGCTGCTGACTGCGCGGAAATGCAACGTTGGCTGGTCTTTTGAT 813
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGCTTCCGGCTCGCACGGTGGTTCGCAAGGAAATCGGAATTTGGG 873
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly 219
Db 874 AATCGGCTTAGCCAGAAAGATCGTGTGCTGGCGCAAGGTGGACTCTACACAGGGA 933
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGCAGCGGAAGAGAGAGCAGT 993
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTCAAGAGTTTCATCTCAGGAGGTGGCCAGACTCCAGGCTCCGGTGCAGCAGATCATG 1053
QY 260 SerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeu 279
Db 1054 TCCCGACCAAGAGGCTGTGTTCATCATGACGGTTTCGATGACCTGGGCTGTGCTCTC 1113
QY 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
Db 1114 AACATGACACAAGCTCTGCAAGACTGGGCTGAGAGAGCCTCCGTTCAACCTCATA 1173
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 GCGAGTGTCTGAGGAAGGTCTGTCTCCCTGAGTCCTTCTGATCGTCAACCGTCAGAGAC 1233
QY 320 ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGCAACAGAAAGCTCAAGTCAGAGTGTGTCTCCCGGTATACCTGTAGTAGAGA 1293
QY 340 IleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
Db 1294 ATCTCCGGGGACAAGAAATCCATTTGCTTCTTGAGCGCGGGATTGGTGACATCAGNAG 1353
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
Db 1354 ACACAAGGGTTGCGTGCATCATCAACACCGTAGCTGCTCGACCGTGCAGTGCAGTGCC 1413
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyLysSerVal 399
Db 1414 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGAGCTGCGAGGACGTGGTGGGGGAGAGCGTC 1473
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
Db 1474 GCCCCTTCAACCAACGCTCACAGGCTGCAGCCGCTTTCGCTTTCATCAGCTCACC 1533
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLysArgPhe 439

Db 1534 CCTCGAGGCGTGTGTCGGCGCTGTCTCAATCTTGGAGGAAAAGAGTTGTCTCTGAAGCGCTTC 1593
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
Db 1594 TGCCTATGGCTGTGGAGGAGGTGTGGAAATAGGAAGTCAGTGTGTGTGATGGTACGACCTC 1653
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1654 ATGCTTCAAGGACTCGGGGAGTCTGAGTCCGTGCTCTGTTCATCATGAAATCCTCTCTC 1713
QY 480 ProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
Db 1714 CCGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTCCAGGACTTCTGT 1773
QY 500 AlaAlaLeuTyrTrpValLeuGluGlyLeuGluProAlaLeuCysProLeuTyr 519
Db 1774 GCGCCTTGTACTACGTGTGAGGGCTGGAAATCGAGCAGCTCTCTGCCCCTCTGTAC 1833
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
Db 1834 GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGAGGCTTCCATATCCACTCGCTT 1893
QY 540 TrpMetLysArgPheLeuPheGlyValSerGluAspValArgArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTCTTGTGCTTGGCCTCGTGAGGAGACGTAAGAGAGCCACTGGAGGTC 1953
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTCCCGTTCCTCGGGGTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTG 2013
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGGTCAGCAGCCTAATGCCACCCAGGAGACACCTCGAGCGCTTCCACTGTCTT 2073
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
Db 2074 TTCAGACTCAAGACAAAGAGTTTGTTCGCTTGCCATTAAACAGCTTCCAAAGAGTGTGG 2133
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCGGATTAAACAGAACCTGGGACTTGTATAGCATCTTCTTCTGCTCCAGCACTGTCCG 2193
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
Db 2194 TATTTGCGGAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGAGATGATCGCTGAG 2253
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
Db 2254 GCATGTCTGTGGTCCCTCTATGGATGCGGGATAAGACCCCTCATTTGAGGAGCAGTGGGA 2313
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCTGCTCCATGCTTGGCACCCACCTCGCGGAGCTGGACCTGGGGCAGCAGC 2373
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCTGACAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCTGCAAG 2433
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGTATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAACACCTCTGGAGA 2493
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 2494 ATGCTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGACCCACCTGGAAGAA 2553
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGTAAGGATGGGTGTGAAGCTTAAACACCCCAAAATGTTTGTGGAGTCTTTG 2613
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGGATTGCTGTGATTGACCCATGCTGTTACTCTGAAGATCTCCCAAAATCCTTACG 2673

QY 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
 DB 2674 ACCTCCCCCAGCTGAATCTCTGAGCCTGCGAGGAACAAGGTGACAGACAGGAGTA 2733
 QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
 DB 2734 ACGCCTCTCAGTGATGCTCTGAGGCTCTCCAGTGGCCCTGCGAGAGCTGATACTGAG 2793
 QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
 DB 2794 GACTGTGGCATCACAGCCAGCGGTGGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 2853
 QY 860 SerLeuThrHisLeuLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
 DB 2854 AGCTTGACACACCTGTGCTCTATCAACACAGCCCTGGGGAACGAGGTGTAAATCTACTG 2913
 QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
 DB 2914 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGCTGATGCTGAATCAGTGGCCAC 2973
 QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
 DB 2974 CTGGACACGGCTGGCTGTGCTTCTTGCACTTGGCTTATGGGTAACCTCATGGCTGACG 3033
 QY 920 HisLeuSerLeuSerMetAsnProValGluAspGlnGlyValLysLeuLeuCysGluVal 939
 DB 3034 CACCTGAGCCCTTAGCATGAAACCCCTGTGGAACACAAATGGCGTGAAGCTTCTGTGCGAGGTC 3093
 QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
 DB 3094 ATGAGAGAACCATCTTGTCACTCCAGGACCTGGAGTTGGTAAAGTGTCACTCACCGCC 3153
 QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
 DB 3154 CGGTGCTGTGAGAGTCTGCTGCTGATCTCGAGGAGCAGACACCTGAAGAGCGCTGGAT 3213
 QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
 DB 3214 CTCAGGCAATGCCCTGGGTGAGCGGTGGGTGCTGCGCTGTCGAGGAGCTGAAGCAA 3273
 QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
 DB 3274 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGAGCTGACTTCTGATGCTGT 3333
 QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
 DB 3334 GAGGCATCTCTCTGGCCCTTCTCTGCAACCGGCATCTGACAGCTTAAACCTGGTGCAG 3393
 QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
 DB 3394 AATACTTTCAGTCCCAAGGAATGATGAAGCTGTGTGGGCTTTGCTGTGCCACGCTCT 3453
 QY 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu 1079
 DB 3454 AACTTACAGATAATGGCTGTGAAATGGCAGTACCTCTGTGCAATAGGAAGCTGCTG 3513
 QY 1080 GluGluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAsp 1099
 DB 3514 GAGGAAGTGCAGCTACTCAAGCCCGAGCTGTAATTCAGCGGTAGTTGGCATCTCTTTGAT 3573
 QY 1100 GluAspAspArgHis 1104
 DB 3574 GAAGATGACCGGTAC 3588

RESULT 6

AAL47135

ID AAL47135 standard; DNA; 5859 BP.

XX AC AAL47135;

XX DT 20-AUG-2002 (first entry)

XX

DE XX Pyrin domain containing protein NALP8/Py12 coding sequence.
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antiporiatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephrotropic; osteopathic; neotropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis; gene; ds.

OS Unidentified.

PN W0200240668-A2.

XX 23-MAY-2002.

XX 30-OCT-2001; 2001WO-EP012545.

XX 15-NOV-2000; 2000DE-01056687.

XX 30-NOV-2000; 2000DE-01059595.

XX (APOT-) APOTEC RES & DEV LTD.

XX Tschoopp J, Martinon F;

XX WPI; 2002-427093/45.

XX P-PSDB; AAO17863.

XX New DNA encoding protein with pyrin domain, useful for treating diseases
 PT involving impaired signal transduction, particularly inflammation, also
 PT proteins and antibodies.

XX Claim 5; Fig 1; 116pp; German.

XX The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a coding sequence of
 CC the invention

SQ Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 5859
 Score: 5724.50 Matches: 1106
 Percent Similarity: 90.53% Conservative: 3
 Best Local Similarity: 90.29% Mismatches: 7
 Query Match: 80.92% Indels: 109
 DB: 6 Gaps: 3

US-10-066-521-6 (1-1344) x AAL47135 (1-5859)

QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTrpGlyLeuGlnTrpCysLeuTrpGlu 20
 DB 1975 ATGGAAGGAGACAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGGTGTCTCTATGAG 2034
 QY 21 LeuAspLysGluCluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40
 DB 2035 CTAGACAAGGAAGAAATTCAGACATTCAGGAATTAATAAGGAAGAAATTCAGGAATCG 2094
 QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
 DB 2095 ACCACATGCTCTATTCCAGATTGGAATCGAGATGCCACGTGGATGTCTGGCAGCTC 2154
 QY 61 LeuLeuHisGluTrpTrpGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
 DB 2155 CTCTTCATGAGTATTATGAGCATCGCTGGCGCTACGTCACGTCCATTAGCATCTTTGAA 2214
 QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaAlaArgAspMetLysLys----- 97


```
QY 780 ArgLeuaspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 4435 AGGCTGGATTCTGTGGATTGACCATGCTGTACCTGGAAGATCTCCCAATCTTTACG 4494
QY 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 4495 ACCTCCCCAGCTGAAATCTCTGAGCTGCGAGNAACAAGGTGACAGACCAGGAGTA 4554
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 4555 ATGCTCTCTCAGTGATGCTTGCAGAGTCTCCAGTGCCTGCGCCCTGCAGAAAGCTGATCTGGAG 4614
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 4615 GACTGTGGCATCACAGCCAGCGGTGCGAGAGTCTGGCCCTCAGCCCTCGTCAGCAACCGG 4674
QY 860 SerLeuThrHisLeuLysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 4675 AGCTTGACACACTGTGCTTATCCACACAGCTTGGGGAACGAGGTGTAAATCTACTG 4734
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 4735 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGTATGCTGAATCAGTGCAC 4794
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 4795 CTGACACGGCTGGCTGGTGGTTCCTTGTGCTTGGCTTATGGGTAACTCATGGCTGACG 4854
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLysGluVal 939
Db 4855 CACCTGAGCTTAGCATGAACCTGTGGAGACATGGCGTGAAGCTTCTGTGGAGGTC 4914
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 4915 ATGAGAGAACCATCTTGTCTATCTCCAGACCTGGAGTTGGTAAAGTGTCTATCTCACCGCC 4974
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 4975 GCGTGTGTGAGAGTGTCTCTGTGATCTCGAGGAGCAGACACCTGAAGAGGCTCGAT 5034
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuLysGluGlyLeuLysGln 999
Db 5035 CTCAGGACATGCTGCTGGGTGACGGTGGGTGCTGCTGCTGCGAGGACTGAAGCAA 5094
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 5095 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGGACTGACTTCTGATTGCTGT 5154
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 5155 GAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCATCTTAAACCTGGTGAG 5214
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuLysSerAlaPheAlaCysProThrSer 1059
Db 5215 AATACTTCACTCCCAAGGAATGATGAGTGTGTTGGGCTTGGCTGTGCCACGCTCT 5274
QY 1060 AsnLeuGlnIleIleGly----- 1065
Db 5275 AACTTACAGATAATGGCAATGACTCTGAAGAAAATGACGTTCTTCGAGAATCTGCTCTA 5334
QY 1065 ----- 1065
Db 5335 GTAGTTTGTAAAGTCACATGTTTCCAGAACCTATCAATGACATTAAGGGAGAACTTA 5394
QY 1065 ----- 1065
Db 5395 CTGTACCTCCCAACCCCTATAACAGCCCGGCACAGAGACTCTGAGGAAGGAATTCAT 5454
QY 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 5455 GGATGGACTGAAGGCTGTGGAAATGGCAGTACCTCTGTGCAATAAAGGAAGCTGCTGGAG 5514
```

```
QY 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 5515 GAAGTGAGCTACTCAAGCCCCGAGTCGTAATGACGGTAGTGTGGCAATCTTTTGATGAA 5574
QY 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 5575 GATGACCGATTGGATCTTCAAAAGCCAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 5634
QY 1112 LeuProGluSerArg 1116
Db 5635 CTCATGCGCATCACAG 5649
RESULT 7
AAL47131
ID AAL47131 standard; DNA; 6939 BP.
XX
AC AAL47131;
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP5/Py8-hs coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
PN W02002040668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-EP012545.
XX
PR 15-NOV-2000; 2000DE-01058687.
PR 30-NOV-2000; 2000DE-01059595.
XX
PA (APOT-) APOTECH RES & DEV LTD.
XX
PI Tschopp J, Martinon F;
XX
DR WPI; 2002-427093/45.
DR P-PSDB; AAO17859.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
XX proteins and antibodies.
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3
```

```
US-10-066-521-6 (1-1344) x AAL47131 (1-6939)
QY 1 MetGluGlyAspLysSerLeuThrPheSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
DB 3055 ATGGAGAGAGCAAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGGTGTCTCTATGAG 3114
QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
DB 3115 CTAGACAAGGAAGAAATTTACAGACATTCAGGAAATTTACTAAAGAAGAAATCTTCAGAATCG 3174
QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
DB 3175 ACCACATGCTCTATTTCCACAGTTTGAATTCAGAAATGCCAAACGTGGAATGTCTGGCACTC 3234
QY 61 LeuLeuHisGluTyrTyrGlyValaserLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
DB 3235 CTCCTTGATGATATATATGGAGCATCGCTGGCCCTGGGCTACGTCCATAGCATCTTTGAA 3294
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
DB 3295 AACATGAACCTTCGAACCTCTCTCGAGAAGGCACGGGATGACATGAAAGACATTCACCA 3354
QY 97 ----- 97
DB 3355 GAAGATCCTGAAAGCAACGATGACTGACCAAGGACCAAGGAAAAAAGTGCCAGGAATT 3414
QY 98 -----IleSer 99
DB 3415 TCACAAGCTGTGCACAAAGATAGTGCCACAGCTGCAGAGACAAAGAACAAAGAAATTTCA 3474
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln 119
DB 3475 CAAGCTATGGAACAAAGAGGTGCCACAGCAGCAGAGACAGAGAACAAGAAATTTCCAA 3534
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyAsp 139
DB 3535 GCTATGGAACAAAGAGGTGCCACAGCAGCAGCAGACAGAGAACAAGGACATGGAGGTGAC 3594
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
DB 3595 ACATGGGACTACAAGATCACTGATGACCAAAATTCGCTGAGGAGGAGATGTACGTCGT 3654
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
DB 3655 AGTTTGAACACACTGCTGCTGACTGGCCGGAATGCAAACTTTGGCTGGTCTTTTGAT 3714
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
DB 3715 TCAGACCGTGGGGCTTCGGGCTCGCACCGGTGTTCTGCACGGAAAGTCAGGAATTTGGG 3774
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly 219
DB 3775 AAATCGGGCTCTAGCCAGAAAGATCGTGTGTGTGGCGCAAGGTGGAACCTTACCCAGGGA 3834
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnAlaArgLysLysSerSer 239
DB 3835 ATGTTCTCTACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3894
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
DB 3895 GTCACAGATTCATCTCCAGGGAGTGGCCAGACTCCCAAGGCTCCGCTGACGGAGATCATG 3954
QY 260 SerArgProGluArgLeuPheIleAspGlyPheAspAspLeuGlySerValLeu 279
DB 3955 TCCCGACAGAAAGCTGTGTGTTCATCATTCAGCGTTTCGATGACCTGGGCTCTGTCTCT 4014
QY 280 AsnAsnAspThrLysLysCysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
DB 4015 AACATGACACAAAGCTCTGCAAGACTGGGCTGGAAGAGAGGCTTCGTTCAACCTCATTA 4074
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
DB 5155 GCATGTCTGTGGTCCCTCTATGATGCGGGATAAGACCCCTCATTTGAGGAGCAGTGGGAA 5214
```

```
DB 4075 CGCAGTCTGCTGAGGAAGGTCTCTCTCCCTGAGTCTCTCTGATCGTCACCGTCAGAGAC 4134
QY 320 ValGlyThrGluLysLeuLysSerSerGluValValSerProArgTyrLeuLeuValArgGly 339
DB 4135 GTGGGCACAGAGAGCTCAAGTCAGAGTCTGTCTCTCCCGTTACCTGTGTATTGAGGA 4194
QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 4195 ATCTCCGGGGAAACAAGAAATCCACTTGTCTCTTGGCGCGGATTTGGTGGACATCAGAA 4254
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
DB 4255 ACAAAAGGGTTGGTGGCATCAACCAACCGTGGCTGCTCGACCAAGTCCAGGTGCC 4314
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399
DB 4315 GCGTGGGCTCTCTCATCTCGCTGGCCCTGACGCTGCAGGACGTGTGTGGGGAGAGGCTC 4374
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
DB 4375 GCCCCTTCAACCAACCGCTCACAGGCTGCACGCGCTTTTGTGTTCATCAGCTCACC 4434
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
DB 4435 CCTCGAGGCGTGTGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTGAAGCGCTTC 4494
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
DB 4495 TGCGTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTGACGTGACGCTC 4554
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 4555 ATGTTCAAGGACTCGGGGAGTCTGAGTCCGTGCTCTGTTTTCACATGAACATCTCTTC 4614
QY 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 4615 CCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGCTCTCCAGGACTTCT 4674
QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
DB 4675 GCGCCTTGTACTACGTGTAGAGGCGCTGGAATCGAGCAGCTCTCTGCCCTCTGTGAC 4734
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 4735 GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGCAGGCTTCCATATCCACTCGCTT 4794
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
DB 4795 TGGATGAAGCGTTTCTTGTGTGGCTCTGTGAGGAGACGCTAAGAGGAGGAGGAGGAG 4854
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 4855 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAAAGCTTCTGCACTGGGTCTCTCTG 4914
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGGTCAGCAGGCTTAATGCCACCCAGGAGACACCTTGGAGCGCTTCCACTGTCTT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACAGCTTCCAAGAAGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCCGATTAACAGAACCTGGGACTTGTATGACATCTTCTTCTGCTTCCAGCACTGTCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTGGGAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGATGATGTCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
DB 5155 GCATGTCTGTGGTCCCTCTATGATGCGGGATAAGACCCCTCATTTGAGGAGCAGTGGGAA 5214
```

```
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 5215 GATTTCGTCTCATGCTTGGCACCCACCACCTCTGGCGAGCTGGACCTGGGGCAGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLeuThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 5275 ATCTTGACAGCGGGCCATGAGACCCCTGTGTCCCAAGCTGAGGCATCCACCTGCAAG 5334
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 5335 ATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGCACCTCTGGAGA 5394
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGlu 759
Db 5395 ATCGTCATGGCCAAACCGTAACCTAAGATCCCTCAACCTTGGAGGACCCACCTGAAGGAA 5454
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 5455 GAGGATGTAAAGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTTGTGGAGCTTTG 5514
QY 780 ArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 5515 AGGCTGGATTCTGTGGATTGACCATGCTCTTACCTGAAGATCTCCCAAAATCCTTACG 5574
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 5575 ACCTCCCCCAGCCCTGAATCTCTCAGCGCTGGCAGAAACAGGTGACAGACCCAGGGAGTA 5634
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 5635 ATGCCTCTCAGTGATGCTTGAGAGTCTCCAGTGGCCCTGCGAGAGCTGATCTGGAG 5694
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 5695 GACTGTGGCATCACAGCACCGGTGGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCCG 5754
QY 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGlyValAsnLeuLeu 879
Db 5755 AGCTTGACACACCTGTGCTATCCAAACACAGCTGGGGACGAAAGGTGTAATCTACTG 5814
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 5815 TGTGATCCATGAGCTTCCCATCTAGTCTGAGAGGCTGATGCTGAATCAGTGGCCAC 5874
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 5875 CTGGACACGGCTGGCTGTGGTTCCTTGCACTTGGCTTATGGGTAACTCATGGCTGACG 5934
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 5935 CACCTGAGCCCTTAGCATGAACCCCTGTGGAAGACAATGGCGTGAAGCTTCTGTGGAGGTC 5994
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 5995 ATGAGAGAACCATCTTGTATCTCCAGGACTTGAGTTGGTAAAGTGTCTATCTCACCGCC 6054
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 6055 CGCTGCTGTGAGAGTCTGTCTGTGTATCTTCGAGGAGCAGACACCTGAAGAGCCTGGAT 6114
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 6115 CTCACGACAATGCCCTGGGGTGGCGGTGGGTGCTGCTGCTGCTGCGAGGAGCTGAAGCAA 6174
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 6175 AAGAACAGTGTCTTGACGAGACTCGGGTGGAGGCATGTGACTGACTTCTGATGTGCTG 6234
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 6235 GAGGCACTCTCTTGGGCCCTTCTCTGCAACCGGCATCTGACCACTGTAAACCTGGTGCAG 6294
```

```
QY 1040 AsnAsnPheSerProLysGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSer 1059
Db 6295 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGGCCTGCCACGTCT 6354
QY 1060 AsnLeuGlnIleIleGly----- 1065
Db 6355 AACTTACAGATANTGGCAATGACTCTGAAGAAAATGACGTTCTTCGAGAATCTGCTCTA 6414
QY 1065 ----- 1065
Db 6415 GTAGTTTTCCTTAAAGTCACTGTTTCCAGAACCTATCAATGACATTAAAGGAGAACTTA 6474
QY 1065 ----- 1065
Db 6475 CTGTACTCTCCCAACCCCTATTAACAGACCCCGGCACAGAGACTCTGAGGAGGAATTCAT 6534
QY 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 6535 GGATGGACTGAAGGCTGTGGAATGGCAGTACCTCTGCAAAATAAGGAAGCTGCTGGAG 6594
QY 1081 GluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 6595 GAAGTGCACTACTCAAGCCCGAGTCGTAATTTGACGCTAGTTGGCATCTTTTGTGAA 6654
QY 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 6655 GATGACCGATTGGATCTTCAAGCCAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 6714
QY 1112 LeuProGluSerArg 1116
Db 6715 CTCATGGCATCACAG 6729
RESULT 8
AAL47140
ID AAL47140 standard; DNA; 6939 BP.
XX AC AAL47140;
XX XX
DT 20-AUG-2002 (first entry)
XX DE Pyrin domain containing protein NALP13/Py17 coding sequence.
XX KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
XX KW antiarteriosclerotic; antiporiatic; antibacterial; virucide;
XX KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
XX KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
XX KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
XX KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX KW osteoarthritis; glomerulonephritis; gene; ds.
XX OS Unidentified.
XX PN WO200240668-A2.
XX XX
XX PD 23-MAY-2002.
XX PF 30-OCT-2001; 2001WO-EP012545.
XX XX
XX PR 15-NOV-2000; 2000DE-01056687.
XX PR 30-NOV-2000; 2000DE-01059595.
XX XX
XX PA (APOT-) APOTECH RES & DEV LTD.
XX PI Tschopp J, Martinon F;
XX XX
XX DR WPI; 2002-427093/45.
XX DR P-PSDB; AAO17868.
XX XX
XX PT New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX PS Claim 5; Fig 1; 116pp; German.
```


XX The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PVD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention

XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3

US-10-066-521-6 (1-1344) x AAL47140 (1-6939)

Qy	1	MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu	20
Db	3055	ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG	3114
Qy	21	LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer	40
Db	3115	CTAGCAAGGAAGAATTTTCAGACATTCAGGAATTTACTAAAGAGAAGAATCTTCAGAAATCG	3174
Qy	41	ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu	60
Db	3175	ACCACATGCTCTATTCCACAGTTTGAATTCAGAAATGCCAAGCTGGAAATGTCTGGCACTC	3234
Qy	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu	80
Db	3235	CTCTTGATGATGATTATGAGAGCATCGCTGGCCCTGGGCTACGTCCATTAGCATCTTTGAA	3294
Qy	81	AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys-----	97
Db	3295	AACATGAACCTGCGAACCTCTCGAGAAAGGCACGGATGACATGAAGAAAGACATTCACCA	3354
Qy	97	-----	97
Db	3355	GAAGATCTCTGAAGCAACGATGACTGACCACGACCAAGCAAGAAAGTCCAGGAATT	3414
Qy	98	-----IleSer	99
Db	3415	TCACAAGCTGTGCAACAGATAGTGCCACAGCTGCAGAGACAAAGAAATTTCA	3474
Qy	100	GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln	119
Db	3475	CAAGCTATGGAACAAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAAGAAATTTCAAA	3534
Qy	120	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGlyAsp	139
Db	3535	GCTATGGACAAAGAGGTGCCACAGCAGCAGCAGACAGACAGACAGACAGACAGACAGACAG	3594
Qy	140	ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg	159
Db	3595	ACATGGGACTCAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCGT	3654
Qy	160	SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp	179
Db	3655	AGTTTTGAAAAACATGCTGCTGACCTGGCCGGAATTCAGAAAGTGGCTGGCTTTTGA	3714
Qy	180	SerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly	199
Db	3715	TCAGACCGGTGGGCTTCCGGCTTCGACCGTGGTTCGACCGAAGATTCAGGAATTTGG	3774
Qy	200	LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGly	219

Db	3775	AAATCGGCTCTAGCCAGAGGATCGTCTGTCTGGCGCGCAAGGTGACCTTACCAGGGA	3834
Qy	220	MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer	239
Db	3835	ATGTTCTCTACGTCTTCTTCCCTCCGTTAGAGAGATGACGCGGAAGAGAGAGCAGT	3894
Qy	240	ValThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet	259
Db	3895	GTCAAGAGTTCTATCTCCAGGAGTGCCAGACTCCCCAGGCTCGGTGACGAGATCATG	3954
Qy	260	SerArgProGluArgLeuLeuPheIleAspGlyPheAspLeuGlySerValLeu	279
Db	3955	TCCCGACCAAGAGGCTGTGTTTCATCATTCAGCGGTTTCGATGACCTGGGCTCTGTCTC	4014
Qy	280	AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle	299
Db	4015	AACATGACAAAGCTCTGCAAGAGCTGGCTGAGNAGAGAGCTCGTTTCACTCATTA	4074
Qy	300	ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp	319
Db	4075	CGCAGTCTGCTGAGGAAGTCT	4134
Qy	320	ValGlyThrGluLysLysSerGluValValSerProArgTyrLeuLeuValArgGly	339
Db	4135	GTGGGCACAGAGAGCTCAAGTCAGAGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4194
Qy	340	IleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLys	359
Db	4195	ATCTCCGGGAAACAAGAAATCCACTTGTCTCTTGTAGCGCGGATTTGTGTAGCATCAGAG	4254
Qy	360	ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro	379
Db	4255	ACACAAGGTTGCTGCGATCATGAACAACCGTGAGTGTCTGCACCAAGTCCAGGTGCC	4314
Qy	380	AlaValGlySerLeuLeuCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal	399
Db	4315	GCCTGGGCTCTCTCATCTGCTGGGCTGTGAGTGCAGACGCTGTGTGGGGAGAGGCTC	4374
Qy	400	AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr	419
Db	4375	GCCTCTTCAACAAAGCTCACAGGCTGCACGCGCTTTTGTGTTCATCAGCTCACC	4434
Qy	420	ProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLysArgPhe	439
Db	4435	CCTCGAGGCTGTGCGCGCTGTCTCAATCTCGAGGAAAGAGTTGTCTGAAAGCGCTC	4494
Qy	440	CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu	459
Db	4495	TGCCGTATGGCTGTGGAGGAGGTGGGAATAGGAAGTCACTGTTTACGCGTACGACCTC	4554
Qy	460	MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu	479
Db	4555	ATGTTTCAAGACTCGGGAGTCTGAGCTCGTCTCTGTTTTCATGAACATCTCTTCTC	4614
Qy	480	ProAspSerHisCysGluGlyTyrTyrThrPheHisLeuSerLeuGlnAspPheCys	499
Db	4615	CCAGACGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTCCAGGACTTCTGT	4674
Qy	500	AlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr	519
Db	4675	GCCTGCTTGTACTACGTGTAGAGGCGCTGGAAATCGAGCAGCTCTCTGCTCTCTCTCT	4734
Qy	520	ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu	539
Db	4735	GTTCAGGAAGCAAGAGGTCATGGAGCTTAAACAGCAGGCTTCCATATATCCACTCCGCT	4794
Qy	540	TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal	559
Db	4795	TGATGAGAGGTTCTTGTGTGGCTCTGTGAGCGAAGACGTAGAGGCCACTCGAGGCTC	4854
Qy	560	LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisIleTrpValSerLeu	579
Db	4855	CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGGCTTCTGACACTGGGTCTCTCTG	4914

```
QY 580 LeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGGTGAGCAGCTTAATGCCACACCCAGAGACACCTTGGACGCTTCCACTGTCTT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTCGAGACTCAAGACAAAGAGTGTTCGCTTGGCATTAAACAGCTTCCAAGAGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCCGATTAACCAAGAACCTTGGACTGTATAGCATCTTCTCTGCCCTCCAGCATGTCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTCGGGAANAATTCGGGTGGATGTCAAGGGATCTTCCCAAGATGAGTCCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
DB 5155 GCATGTCTGTGGTCCCTCTATGATGGGATGAAGACCTCATTTGAGGAGCAGTGGGAA 5214
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 5215 GATTTCGTCTCATGCTTGGCACCACCCACACCTGGCGGACGTGGACCTGGGCGAGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
DB 5275 ATCTGACAGACGGGGGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCTGCAAG 5334
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
DB 5335 ATACAGACCTTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGACGACCTCTGGAGA 5394
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
DB 5395 ATCGTCATGGCAACCGCTAAGATCCCTCAACTTGGGAGGACCCACCTGAAGGAA 5454
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisIleProLysCysLeuLeuGluSerLeu 779
DB 5455 GAGGATGTAGAGTGGGTGTGAAGCCTTTAAACACCCAAATGTTTGTGGAGTCTTTG 5514
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
DB 5515 AGGCTGGATTGCTGTGGATTGACCCATGCTTACCTGAAGATCTCCCAAACTCTTACG 5574
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
DB 5575 ACCTCCCCCAGCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCCAGGAGTA 5634
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
DB 5635 ATGCCTCTCAGTGAGTCTGAGAGTCTCCAGTGCCTGCGCCCTGCGAAGCTGTACTGGAG 5694
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
DB 5695 GACTGTGGCATCACAGCCAGGGTTGCCAGAGTCTGGCCCTCAGCCCTCGTCAGCAACCGG 5754
QY 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
DB 5755 AGCTTGACACACCTGTGCTATCCAAACAACAGCCCTGGGGAAACGAAGGTGTAAATCTACTG 5814
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
DB 5815 TGTGCATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGAGTGTGAATCATGTGCCAC 5874
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
DB 5875 CTGGACAGGCTGCGTGGTGTTCCTTTCGACTTCCGCTTATGGGTAACTCATGGCTGACG 5934
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
DB 5935 CACCTGAGCCTTAGCATGAACCTGTGGAAAGACAATGGCGTGAAGCTTCTGTGCGAGGTC 5994
```

```
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
DB 5995 ATGAGAGAACCATCTTTGTATCTCCAGGACCTGGAGTTGTAAGTGTATCTATCCCGCC 6054
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
DB 6055 GCGTGTGTGAGAGTCTGTCTCTGTGTCTGAGGAGCAGACACCTGAAGACCTGGAT 6114
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
DB 6115 CTCACGACCAATCCCTGGGTGACGGTGGGGTGTGTCACCTGTGCGAGGACCTGAAGCAA 6174
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
DB 6175 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGT 6234
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
DB 6235 GAGGCACCTCTCTTGGCCCTTTCCTGCAACCGGCATCTGACCAGTCTAAACCTGGTGCAG 6294
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
DB 6295 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGCCTGTCCACGTCT 6354
QY 1060 AsnLeuGlnIleGly----- 1065
DB 6355 AACTTACAGATAATTGGCAATGACTCTGAAGAAATGACGTTCTTCGAGAATCTGCTCTA 6414
QY 1065 ----- 1065
DB 6415 GTAGTTTTCGTTAAAGTCACTGTTTCCAAAGACCTATCAATGACATTAAAGGAGAACTTA 6474
QY 1065 ----- 1065
DB 6475 CTGTACTCTCCCAACCCCTATAACACGACCCGACAGAGACTCTCGAGGAAGAAATTCAT 6534
QY 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
DB 6535 GAGTGAAGTCAAGAGCTGTGGAATGCGAGTACCTCTGCNAATATAGGAAGCTGTGGAG 6594
QY 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
DB 6595 GAAGTCAGCTACTCAAGCCCGAGTCGTAATTGACGGTAGTTGGCATCTTTTGTGATGA 6654
QY 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
DB 6655 GATGACCGATTGGATCTTCAAAGCCAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 6714
QY 1112 LeuProGluSerArg 1116
DB 6715 CTCATGGCATCAG 6729
RESULT 9
ADA45220
ID ADA45220 standard; DNA; 3830 BP.
XX ADA45220;
XX
XX 20-NOV-2003 (first entry)
XX
DE Human MATER splice variant DNA SEQ ID 3.
XX human; MATER; maternal antigen that embryos require; NTP-ase;
XX nucleoside triphosphate-ase; apoptosis; antifertility; contraceptive;
XX antiinflammatory; immunosuppressive; gynaecological; endometriosis;
XX ovarian dysfunction; autoimmune premature ovarian failure;
XX autoimmune disease; contraceptive; autoantibody; ds; gene.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..3432
XX /*tag= a
FT
```

/product= "splice variant of MATER"

FT EP1285964-A2.
 PN 26-FEB-2003.
 XX 12-JUL-2002; 2002EP-00090246.
 PD 10-AUG-2001; 2001DE-01039874.
 XX (SCHD) SCHERING AG.
 PA Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
 PI WPI; 2003-302814/30.
 DR P-PSDB; ADA45221.
 XX
 PT New nucleic acid for human maternal antigen that embryos require protein,
 PT useful e.g. in diagnosis and treatment of female infertility.
 XX
 PS Claim 1; Page 17-18; 31pp; German.
 XX
 CC This invention describes a novel human MATER (maternal antigen that
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
 CC associated with apoptosis, a defect in MATER activity causes growth
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
 CC MATER gene is found predominantly in ovary, testis and placenta. The
 CC products of the invention have anti-infertility, contraceptive,
 CC anti-inflammatory, immunosuppressive and gynaecological activity.
 CC Effectors of the MATER polypeptide are used to treat infertility.
 CC associated with endometriosis and also ovarian dysfunction, autoimmune
 CC premature ovarian failure, inflammation, autoimmune diseases and female
 CC infertility, and as contraceptives. Measurements of autoantibodies
 CC against MATER in body samples or MATER protein or mRNA in eggs is useful
 CC for diagnosis of female infertility, particularly by detecting mutations
 CC in the MATER gene with a DNA chip. This sequence represents a splice
 CC variant of human MATER in which exon 4 of the usual form is absent.
 XX
 SQ Sequence 3830 BP; 943 A; 946 C; 1041 G; 900 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: 0 Length: 3830
 Score: 5683.50 Matches: 1087
 Percent Similarity: 95.87% Conservative: 5
 Best Local Similarity: 95.43% Mismatches: 12
 Query Match: 80.34% Indels: 35
 DB: 7 Gaps: 2

US-10-066-521-6 (1-1344) x ADA45220 (1-3830)

QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
 DB 1 ATGGAAGAGAGACAAATCGCTCCACCTTTCCAGCTACGGCTGCAATGGTGCTCTATGAG 60
 QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
 DB 61 CTAGACAGGAAGAATTTCAGACATTCAGGAATTAAGAGAAATCTTCAGAAATCG 120
 QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
 DB 121 ACCACATGCTCTATTCCACAGTTTGAATTCGAGAAATGCCAACGTGGAATGTCTGGCACTC 180
 QY 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
 DB 181 CTCCTTGCATGAGTATTATGAGCATCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 240
 QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGln 100
 DB 241 AACATGAACCTCGGAACCTCTCGGAGAGGACCGGATGACATGAAAAATTCACCAGAA 300
 QY 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln--- 115
 DB 301 GATCTGGAAGCAACGATGACTGACCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360

QY 115 ----- 115
 DB 361 TATGGCATGACTAAGCTTATCTTTGGGGGTGCTGTGACATCTCTGACTCGAATAATAAACAC 420
 QY 116 -----GluIleSerGlnAlaMetGluGlnGluGly 125
 DB 421 AAGTATGTTGGAATTCATTCTTTTTCAGAAATTTTTCACAACTATGGAACAAAGAGGT 480
 QY 126 AlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyVAspThrTrpAspTyrLysSer 145
 DB 481 GCCACAGCAGCAGACAGACAGAAAGACATGAGGTGACACATGGGACTACACAGT 540
 QY 146 HisValMetThrLysPheAlaGluGluGluAspValArgSerPheGluAsnThrAla 165
 DB 541 CACGTGATGACCAAAATTCGCTGAGGAGGAGTGTACGTCTAGTTTGTGAAACACCTGCT 600
 QY 166 AlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPhe 185
 DB 601 GCTGACTGGCCCGGAAATGCAAACTGTTGGCTGGTCTTTTGAATTCAGACCCGTGGGGCTTC 660
 QY 186 ArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 205
 DB 661 CGGCTCGCACGGTGTCTCTGACGGAAGTTCAGAAATTCGGAATTCGGCTCTAGCCAGA 720
 QY 206 ArgIleValLeuCysTrpAlaGlnGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhe 225
 DB 721 AGGATCGTGTGCTGGCGGCAAGGTGGACTCTACCAGGGAATGTTCTCTACGTCTTC 780
 QY 226 PheLeuProValArgGluMetGlnArgLysGlyLysSerValThrGluPheIleSer 245
 DB 781 TTCTCTCCCTGTAGAGAGATGACGCGGAAGAGAGAGAGAGTGTACAGAGTTCATCTCC 840
 QY 246 ArgGlnTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeu 265
 DB 841 AGGAGTGGCGCAGACTCCAGGCTCCGGTCACGAGATCATGTCCCGACCAAGAGCTG 900
 QY 266 LeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeu 285
 DB 901 TTGTTTCATCATTCAGCGGTTTCGATGACCTGGGCTCTGTCTCAACATGACACAAAGCTC 960
 QY 286 CysLysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeuArgLys 305
 DB 961 TGCAAGACTGGGCTGAGAAAGACGCTCCGTTCCACCTCATACGACGTCTGCTGAGGAAG 1020
 QY 306 ValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeu 325
 DB 1021 GTCTGTCTCTCTGAGTCTTCTCTGATCGTCAGACGCTGGGACAGAGAGCTC 1080
 QY 326 LysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArg 345
 DB 1081 AAGTCAGAGTCTGTCTCCCGTTACTCTGTAGTAGAGGAATCTCCGGGGAACAAAGA 1140
 QY 346 IleHisIleLeuLeuGluArgGlyIleGlyGlyHisGlnLysThrGlnGlyLeuArgAla 365
 DB 1141 ATCCACTTGTCTCTTGGCGGGGATTTGGTGAGCATCAGAAAGACACAAAGGTTGCGTGC 1200
 QY 366 IleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIle 385
 DB 1201 ATCATGAACACCGTGAGTGTCTCGACAGTCCAGGTCGCCGCTGGGCTCTCTCATC 1260
 QY 386 CysValAlaLeuGlnLeuGlnAspValValGlyGlySerValAlaProPheAsnGlnThr 405
 DB 1261 TGCCTGGCCCTGACGCTGCAGAGGCTGTGGGGGAGAGCGTCGCCCTCTCAACCAACG 1320
 QY 406 LeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArg 425
 DB 1321 CTCACAGGCTGCACCGCGCTTTTGTGTTTCATCAGCTCACCCCTCAGGCGGTGGTCCGG 1380
 QY 426 ArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGlu 445
 DB 1381 CGCTGTCTCAATCTGGAGGAAGAGTGTCTCTGAAAGCGCTCTCTGCCGTATGGCTGTGGAG 1440

QY 446 GlyValTprAsnAlaGlySerValPheAspGlyAspLeuMetValGlnGlyLeuGly 465
DB 1441 GGAGTGTGAATAGGAAGTCAAGTGTGGAGCGTGAACCTCAATGGTTCAAGGACTCGG 1500
QY 466 GluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGlu 485
DB 1501 GAGTCTGAGCTCCGTGCTCTGTTTCAATGAACATCTCTTCCAGACAGCCACTGTGAG 1560
QY 486 GluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaLeuTyrTyrVal 505
DB 1561 GAGTACTACACCTTCTCCACTCAGTCTCCAGGACTTCTGTGCGGCTGTGTACTAGGTG 1620
QY 506 LeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLeuYsThrLysArg 525
DB 1621 TTAGAGGCGCTTGGAAATCGAGCCAGCTCTCTGCGCCCTCTGTACGTTGAGAAGCAAGAGG 1680
QY 526 SerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTyrMetLysArgPheLeu 545
DB 1681 TCCATGGAGCTTAAACAGCGAGCTTCCATATCCACTCGCTTTCGATGAAGCGTTCTTG 1740
QY 546 PheGlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProVal 565
DB 1741 TTTGGCCCTCGTGAAGCAAGAGTAAAGAGGCGCACTGGAGGTCTCTGCTGGGCTGTCCGGTT 1800
QY 566 ProLeuGlyValLysGlnLysLeuLeuHisTyrPValSerLeuLeuGlyGlnGlnProAsn 585
DB 1801 CCCCTGGGGGTGAAGCAGAGCTTCTGACCTGGGTCTCTCTGTTGGGTGACAGCCTAAT 1860
QY 586 AlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLys 605
DB 1861 GCCACCACCCAGGAGACACCTCGAGCGCTTCCACTGCTCTTTTCGAGACTCAAGACAAA 1920
QY 606 GluPheValArgLeuAlaLeuAsnSerPheGlnGluValTprLeuProIleAsnGlnAsn 625
DB 1921 GAGTTTGTTCCTTGGCATTAAACAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAAC 1980
QY 626 LeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArg 645
DB 1981 CTGACTTGATAGCATCTTCTTCTGCTTCAGCAGCTGCTCGTATTTGCGGAAATTCGG 2040
QY 646 ValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValPro 665
DB 2041 GTGGATGTCAAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGCAATGTCCTGTGTCCT 2100
QY 666 LeuTprMetArgAspLysThrLeuIleGluGlnTprGluAspPheCysSerMetLeu 685
DB 2101 CTATGGATGCGGGATAAGACCTCATTTGAGGAGCAGTGGGAAGATTCTTGTCTCCATGCT 2160
QY 686 GlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAla 705
DB 2161 GGCACCCACCCACACTCGCGCAGCTGGACCTGGGCAGCAGCATCTGACAGAGCGGGCC 2220
QY 706 MetLysThrLeuCysAlaLeuLysLeuArgHisProThrCysLysIleGlnThrLeuMetPhe 725
DB 2221 ATGAAGACCTGTGTGCCAAGCTGAGGCATCCCACTTGCAGATACAGACCTCATGTTT 2280
QY 726 ArgAsnAlaGlnIleThrProGlyValGlnHisLeuTprArgIleValMetAlaAsnArg 745
DB 2281 AGAAATGCACAGATTACCCCTGGTGTGCAGCACCTCTGGAGAAATCGTATGCGCCAAACCGT 2340
QY 746 AsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAla 765
DB 2341 AACCTAGATCCCTCACTTGGGAGGCACCCACTTGAAGGAAGAGGATGTAAAGATGGCG 2400
QY 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGly 785
DB 2401 TGTGAAGCCTTAAACACCCAAAATGTTTGTGGAGTCTTTGAGGCTTGGATTGTGTGGA 2460
QY 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
DB 2461 TTGACCCATGCTGTGTTACTTGAAGATCTCCAAAATCTCTTACGACTCCCCAGCCTGAAA 2520
QY 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825

DB 2521 TCTCTGAGCTGCGAGAAACAGGTGACAGACCGGAGTAATGCTCTCACTGATGCC 2580
QY 826 LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAla 845
DB 2581 TTGAGAGTCTCCCAAGTGGCCCTGCAAGAGCTGATACTGGAGGACTGTGGCATCACAGCC 2640
QY 846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
DB 2641 ACCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGC 2700
QY 866 LeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeu 885
DB 2701 CTATCCAAACACAGCTTGGGAAACGAAGGTGTAATCTACTGTGTATCCATGAGGCTT 2760
QY 886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
DB 2761 CCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCCACCTGGACACGGCTGGCTGT 2820
QY 906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTprLeuThrHisLeuSerLeuSerMet 925
DB 2821 GGTTTTCTTGCACCTTGGCTTATGGGTAACTCATGCTGACGACCTGAGCCCTTAGCATG 2880
QY 926 AsnProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCys 945
DB 2881 AACCTGTGGGAAGACATGGCGTGAAGCTTCTGTGCGAGTCAATGAGAGACCATTTGT 2940
QY 946 HisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeu 965
DB 2941 CATCTCCAGACCTGGAGTTGGTAAAGTGTCACTCATCCGCGCGTGTGTGAGAGTCTG 3000
QY 966 SerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeu 985
DB 3001 TCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCTTGTATCTCAGGACCAATGCCCTG 3060
QY 986 GlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThr 1005
DB 3061 GGTGACGGTGGGTTTGTGCACTGTGGAGGACTGAGCAAAAAGAACAGTGTCTTGACG 3120
QY 1006 ArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAla 1025
DB 3121 AGACTCGGGTTGAAGCATGTGGACTGACTTCTGATTGCTGTGAGGCATCTCTCTGGCC 3180
QY 1026 LeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLys 1045
DB 3181 CTTTCTTCGAACCGGCATCTGACAGTCTAAACCTGGTGCAGAAATACTTCAGTCCCAAA 3240
QY 1046 GlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeGly 1065
DB 3241 GGAATGATGAGCTGTGTGGCCTTGTGCTGCCACGCTTAACCTTACAGATAATTGGG 3300
QY 1066 LeuTprLysTprGlnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLeu 1085
DB 3301 CTGTGAAAATGGCAGTACCCCTGTGCAAAATAAGAAAGCTGTGGAGGAAGTGCAGCTACTC 3360
QY 1086 LysProArgValIleLeuAspGlySerTprHisSerPheAspGluAspArgHis 1104
DB 3361 AAGCCCGAGTGTAAATGACGGTAGTTGGCATTTTGTGATGAAGATGACCGGTAC 3417
RESULT 10
ABX97181
ID ABX97181 standard; cDNA; 3226 BP.
XX
AC ABX97181;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV125b cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX

QY 250 pSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAs 270
Db 593 CTCACAGGCTCCGGTGA CGGAGATCATGTCCCGACCAAGAAAGGCTGTGTTCATCATTTGA 652
QY 270 pGlyPheAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAl 290
Db 653 CGGTTTCATGACCTGGGCTCTGTCTCTCAACATGACACAAGCTCTGCAAGACTGGGC 712
QY 290 aGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGl 310
Db 713 TGAGAAGCAGCGCTCCGTTACCCCTCATACGCACTGTCTGAGGAAGGTCTCTGTCTCCCTGA 772
QY 310 userPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValVa 330
Db 773 GTCCCTTCTGTATCGTCAACCGTCAGAGCTGGGCACAGAGAGCTCAAGTCAGAGGTCGT 832
QY 330 lSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLe 350
Db 833 GTCTCCCGTTACCTGTATTAGAGGAATCTCCGGGGACAAGAAATCCACTTGTCTCCT 892
QY 350 uGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAr 370
Db 893 TGAGCGCGGATTTGGTGAGCATCAGAAGACACAAGGGTTGGTGGCATCATGAACAACCG 952
QY 370 gGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGl 390
Db 953 TGAGTGTCTCGACCACTGCCAGGTGCCCGCTGGGCTCTCTCATCTGGCGTGGCCCTGCA 1012
QY 390 nLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHi 410
Db 1013 CCTGCAGGACCTGTGTGGGGAGAGGCTGCCCCCTTCAACCAACCGCTCACAGGCTTGCA 1072
QY 410 -sAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLe 430
Db 1073 CGCCGCTTTTGTGTTTCATCAGCTCACCCCTCGAGGGGTGGTCCGGCGCTGTCTCAATCT 1132
QY 430 uGluLysArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnAr 450
Db 1133 GGAGAAAGAGTGTCTCGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGGGAATAG 1192
QY 450 GlySerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuAr 470
Db 1193 GAAGTCAGTGTTCAGCGTGACGACCTCATGTTCAGGACTCGGGAGTCTGAGCTCCG 1252
QY 470 gAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTrpThrPh 490
Db 1253 TGCTCTGTTTCATGAACATCTCTCCACAGACGACCTGTGTGAGGAGTACTACACCTT 1312
QY 490 ePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGl 510
Db 1313 CTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCTTGTACTACGTGTAGAGGCGCTGGA 1372
QY 510 uIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLy 530
Db 1373 AATCGAGCCAGCTCTCTGCCCTCTGTAGCTTGTAGACACAAGAGGTCCATGGAGCTTAA 1432
QY 530 sGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSe 550
Db 1433 ACAGCAGGCTTCATATCCACTCGCTTTGGATGAAGGTTCTTGTGTGGCCTCGTGAG 1492
QY 550 rGluAspValArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLy 570
Db 1493 CGAAGACGTAAAGAGGCGCACTGGAGGTCTGTGGGCTGTCCGCTTCCCTGGGGGTGAA 1552
QY 570 sGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGl 590
Db 1553 GCAGAGGCTTCTGCACTGGGTCTCTCTGTGTGGGTTCAGAGGCTTAATGCCACCCAGG 1612
QY 590 yAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLe 610
Db 1613 AGACACCTCGACGCTTCCACTGTCTTTTCAGACTCAAGACAAGAGTTTGTGTGCTT 1672

QY 610 uAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAl 630
Db 1673 GGCAATTAACAGCTTCCAAGAGTGTGGCTTCGATTAAACAGAACTGAGACTGTATAGC 1732
QY 630 aSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGl 650
Db 1733 ATCTTCTCTCTCCAGCACTGTCCGATTATTGCGGAAAATTTCGGGTGATGTCAAAGG 1792
QY 650 yIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMetArgAs 670
Db 1793 GATCTTCCCAAGAGATGAGTCGCTGAGGCATGTCTCTGTGGTCCCTCTATGGATGCGGA 1852
QY 670 pLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHi 690
Db 1853 TAAGACCTCATTTGAGAGAGAGTGGGAAGATTTCCTCATCTGTGGCACCCACCACCA 1912
QY 690 sLeuArgGlnLeuAspLeuGlySerIleLeuThrGluArgAlaMetLysThrLeuCy 710
Db 1913 CCTCGGCGAGCTGGACCTGGGACAGCATCTCTGACAGAGCGGCCCATGAAGACCTGTG 1972
QY 710 sAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIl 730
Db 1973 TGCCAAAGCTGAGCATCCCACTGCAAGATACAGACCTGATGTTTAGAAATGCACAGAT 2032
QY 730 eThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLe 750
Db 2033 TACCCCTGGTGTGCAGACCTCTGGAGAAATCGTCATGGCCCAACCGTAACCTAAGATCCCT 2092
QY 750 uAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLy 770
Db 2093 CAACTTGGGAGGCAACCCACTGAAGGAGAGGATGTAAGATGCGCTGTGAACCTTAA 2152
QY 770 sHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCy 790
Db 2153 ACACCAAAATGTTCTTGGAGTCTTTGAGGCTGGATTGCTGTGGATTGACCATGCCTG 2212
QY 790 sTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAl 810
Db 2213 TTACCTGAGAGATCTCCAAATCTTACGACCTCCCGACCTGAAATCTCTGAGCCTGGC 2272
QY 810 aGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGl 830
Db 2273 AGGAAACAAGTGCACAGACCGAGGTAATGCTCTCAGTGATGCTTGCAGAGCTCTCCA 2332
QY 830 nCysAlaLeuGlnLysLeuLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSe 850
Db 2333 GTGCGCCTCGCAGAGCTGATACTGAGGAGCTGTGGCATCACAGCCACGCGTTGCCAGAG 2392
QY 850 rLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSe 870
Db 2393 TCTGGCCTCAGCCTCGTCAGCAACCGAGCTTGACACACCTGTGCTATGCCAACACAG 2452
QY 870 rLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLe 890
Db 2453 CCTGGGGAAAGAGGTGTAATCTACTGTGTGATCCATGAGGCTTCCCACCTGTAGTCT 2512
QY 890 uGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLe 910
Db 2513 GCAGAGGCTGATGCTGAAATCAGTGCCACCTGGACACGGCTGGCTGTGGTTTTCTTGACT 2572
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 2573 TGGCCTTATGGGTAACTCATGGCTGACGACCTTGAGCCTTAGCATGAACCTGTGGAAGA 2632
QY 930 pAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 2633 CAATGGCGTGAACCTTCTGTGCGAGGTCTATGAGAGAACCATCTTGTCACTCCAGGACCT 2692
QY 950 uGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 2693 GGAGTTGGTAAAGTGTCACTCACCCCGCGTGTGTGAGAGTGTCTGTCTGTGTGATCTC 2752
QY 970 rArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyGlyVa 990


```

Db 2753 GAGGAGCAGACACCTGAGAGCCTGGATCTCACGACCAATGCCCTGGGTGACGGTGGGGT 2812
Qy 990 lAlaAlaLeuCySGLuGLyLeuLYsGLnLYsAsnSerValLeuThrArgLeuGLyLeuLY 1010
Db 2813 TCCTGCATCTGCGAGGAGTGAAGCAAAAGACAGTGTCTGACGACACTCGGGTTGAA 2872
Qy 1010 sAlaCySGLyLeuThrSerAspCySGLuAlaLeuSerLeuAlaLeuSerCySAsnAr 1030
Db 2873 GGCATGTGAGCTGACTCTGATTGCTGTGAGGCACTCTCCTTGGGCCCTTCTCTGCAACCG 2932
Qy 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLYsGLyMetMetLYsLe 1050
Db 2933 GCATCTGACCATCTAAACCTGGTGCAGATAACTTCAGTCCCAAGGAAGATGATGAAGCT 2992
Qy 1050 uCySserAlaPheAlaCySProThrSerAsnLeuGlnlleGlyLeuTpyLYsTrpGl 1070
Db 2993 GGTTCGGCCCTTGCTGTCCACGCTTAACCTACAGATAATTGGGCTGTGGAAATGGCA 3052
Qy 1070 nTyrProValGlnlleArgLYsLeuGLuValGlnLeuLeuLYsProArgValVa 1090
Db 3053 GTACCCCTGTGCAATAAGGAAGCTGCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGGT 3112
Qy 1090 lTleAspGLySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3113 AATTGACGGTAGTGGCATCTTTTGTGATGAAGATGACCGGTAC 3155

RESULT 11
ABK48611
ID ABK48611 standard; cDNA; 3447 BP.
XX
AC ABK48611;
DT 13-AUG-2002 (first entry)
DE Mouse WATER cDNA.
XX
KW Mouse; gene; ss; contraceptive; antiinfertility; WATER;
KW maternal antigen that embryos require; WATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; gene therapy; fertility.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT 9..3344
FT CDS /tag= a
FT /product= "Mouse WATER"
XX
XX WO200232955-A1.
XX
XX 25-APR-2002.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX 18-OCT-2000; 2000US-0241510P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2002-454595/48.
XX
XX P-PSDB; AAU79525.
XX
XX New isolated human Maternal Antigen That Embryos Require protein and
XX polynucleotide, useful in diagnosing or treating fertility or reduced
XX fertility, or as a contraceptive.
XX
XX Example 5; Page 72-74; 93pp; English.
XX
XX The invention discloses an isolated human MATER (Maternal Antigen That
XX Embryos Require) protein, which can complement a WATER null phenotype in
XX which zygotes arising from the oocyte do not progress beyond the two-cell

```

```

CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene. The MATER protein and
CC polynucleotide, by gene therapy, are useful in diagnosing or treating
CC fertility and reduced fertility. In particular, the MATER protein is
CC useful as a contraceptive agent, or for influencing (either inhibiting or
CC enhancing) fertility and can be used to detect a predisposition to
CC infertility or reduced fertility, or for presymptomatic screening of an
CC individual for infertility/reduced fertility. The protein and
CC polynucleotide are also useful for detecting an excess or deficiency, or
CC genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the mouse MATER cDNA
XX
SQ Sequence 3447 BP; 1000 A; 783 C; 843 G; 821 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 2.94e-204 Length: 3447
Score: 2688.00 Matches: 554
Percent Similarity: 66.49% Conservative: 186
Best Local Similarity: 49.78% Mismatches: 293
Query Match: 38.00% Indels: 81
DB: 6 Gaps: 11

```

US-10-066-521-6 (1-1344) x ABK48611 (1-3447)

```

Qy 15 GlnTrpCysLeuTyrGluLeuAspLYsGluGluPheGlnThrPheLYsGluLeuLYs 34
Db 163 CAATGGGTCTCCCA-GAAAAAGACAGTAAGCAATCTTGAAGACACCGATTTGGAAGAG 221
Qy 35 LysLYsSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54
Db 222 GACAGAAAGTCAGAAAGCACA----- 242
Qy 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
Qy 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLYsAlaArgAspAsp 94
Db 243 ---ATGTCTCTTCAGAAATATGTCAGTAGAGCAATCTCTGAAAGACAGTGGATCAGAAAGA 299
Qy 95 MetLYsLYsIleSerGln----- 100
Db 300 GTGGACACGGCGTCAGAAAGAAAAATGACTTCTCCAGAAAAACGACAGTAAATCAATCCAG 359
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 AAAGACCAAGGACCAGAGCAGAGCAGACATCAGAAACCTTACAATCTAAGGAAGAAGAT 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GAAGTGACAGAGCA-----GATAAAGAT 443
Qy 136 HisGlyGlyAspThrTrpAspTyrLYsSerHisValMetThrLYsPheAlaGluGluGlu 155
Db 444 AATGGAGTGTACTTACAGACTACAGGCCCAATGTGATTGCTAAGTTGCGACACAGGTGTG 503
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCGACAGATGAAATTTATGTCT 542
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLYs 195
Db 543 GATCCTTTTAAACCATACAGAAACCTTCCAGCTCACACCATATATCCATACAGGAAGA 602
Qy 196 SerGlyIleGlyLYsSerAlaLeuAlaArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 603 CCAGAGTGTGGAAAGTCACTTTGGCCAGAAATATTGTTCTTGGCTGGGCACAGGGTAA 662
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLYs 235

```

Db 663 CTCITCCAAAAATG---TCCTTTGTCTATCTTCTCTCTAGAGAAATAAAGTGGACA 719
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 720 GAGAAAGACAGTTGGCACACGCTGATTGCTAAGGAGTGTCAGAGCTCCTGGGATCTAGTG 779
Qy 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeu 275
Db 780 ACAAGATCATGTCACACACAGAGACTCTGTGTCATAGATGGCTTGCGATGATG 839
Qy 276 GlySerValLeu---AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 840 GACTCTGTCTCCACATGATGATGACACTATCCAGAGACTGGAAGGATGAACAGCCC 899
Qy 295 ProPheThrIleLeuArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 900 ATATACATCCCTGATGATCAGCCCTCTGAGGAAGGCTCTCTTACCTCAGTCCCTTCTCATC 959
Qy 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
Db 960 ATTACCACAGAAACACAGGCTTAGAAAACTCAAGTCAATGGTTGTGTCCTCCCTCTAT 1019
Qy 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 1020 ATACTGGTTGAAGGACTGTCTGCATCAAGGAGATCTCAGCTGTGTCCTCGAGAACATCTCC 1079
Qy 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
Db 1080 AATGAGTCTGATAGAAATACAAAGTCTTCCATTCTCTGATAGAAAATCACCAGCTGTTTGAC 1139
Qy 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuLeuAspVal 394
Db 1140 CAATGCCAGGCCCTCTGTGTGCTCCTGTGCTGTGAGGCTCTACAGCTACAGAGAAA 1199
Qy 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
Db 1200 CTGGMAAGAGATGCACCTACCCTGCCAGACTCTCACCGGTTTGTATGCCACGTTGGTG 1259
Qy 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgVal 434
Db 1260 TTTTACCAGCTCACCTTGAAAGGCCCTTCCAGAGCGCTCTCAGCTCAGGAAGAACAGATT 1319
Qy 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
Db 1320 ACTCTAGTGGTTGTGCTGATGATGCGAGCTGAAGAGTGTGGACCATGAGTCGGTGTTC 1379
Qy 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
Db 1380 TATGATGATGACCTGAAGAACTATAGCTTAAAGGAGTCTGAGATCTTGGCCCTCTTTTAC 1439
Qy 475 MetAsnIleLeuLeuProAspSerHis---CysGluGluTyrThrPhePheHisLeu 493
Db 1440 ATGAACATCCTTCTCCAGGTTGGCCACAAAGTGAAGCAGTGTATGTTTCTCCACCTC 1499
Qy 494 SerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeu---GluIleGlu 512
Db 1500 AGCCTGAGGATTTCTTGTGCTGCTTATATATGTTTGAAGGCTGGAGGATGGAAT 1559
Qy 513 ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla 532
Db 1560 CAGCATTTTTCG-----TTTCATTGAAACCCAAAGGAGCATCATGGAGGTGAGAGAACT 1613
Qy 533 GlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAsp 552
Db 1614 GAC---GACACTCGCCCTCTCGGATGAAGCGTTTCTTATTTGGCCCTCATGAACAAGGAT 1670
Qy 553 ValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLys 572
Db 1671 ATCTTGAAAGACTCTGGAGGTTCTGTTGAATATCCCGTGATTCCACTGTTGAGAGAG 1730
Qy 573 LeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThr 592
Db 1731 CTCCAAACACTGGGTCTCTCTGATAGCTCAGCAGGTCAATGGCACCCAGCCCAATGGACACC 1790

Qy 593 LeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeu 612
Db 1791 CTGGATGCCCTTCTATTTGTTCTATTTGAGTCTCAGGATGAAGAGTTTGTGGGGGGCTCTC 1850
Qy 613 AsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSer 632
Db 1851 AAACGGCTTCCAGAGAGTGGCTGTGATTAACAGAGATGGACTTTGAAGTCTCTTCC 1910
Qy 633 PheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePhe 652
Db 1911 TACTGTCTCAAGCACTGTGAGAACTTGAAGGCAATCCGGGTGGATATCAGAGACCTCCTC 1970
Qy 653 ProArgAspGluSerAlaGluAlaCysProValValProLeuTrp---MetArgAspLys 671
Db 1971 TCGGTAGATAAATCTCTCGAGCTGTCCCTGTTGTTACTGTCCAGGAGACACAATGTAAG 2030
Qy 672 ThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeu 691
Db 2031 CCCTCTCATGAGTGTGGGAAAACCTTCTGCTCTGTCTTGGCAGCCTCCGGAACTTG 2090
Qy 692 ArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAla 711
Db 2091 AAGGAGCTGGACTTGGCGACAGCATCTCTGAGTCAACGGGCCCATGAAGATCTGTGCCTC 2150
Qy 712 LysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThr 731
Db 2151 GAGCTCGGAATCAGTCTGCGAATACAGAACTGACGTTTAAAGATGCGAGAGTAGTG 2210
Qy 732 ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsn 751
Db 2211 TCTGGCTGAAACANUCTTGGAAAGCTCCTTTTAGCAATCAAAACTTAAAGTACCTCAAT 2270
Qy 752 LeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHis 771
Db 2271 CTAGGGAACACTCCCATGAAGATGATGATGAGTCTAGCTCGAAGCGCTGGAACAT 2330
Qy 772 ProLysCysLeuLeuSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyr 791
Db 2331 CCAAGTGTCTCGGAGACTCTGAGGTTGGATTCCTGTGAGTTAACCATCATTTGTTAT 2390
Qy 792 LeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGly 811
Db 2391 GAGATGATCTCCAGCTTCTTATTTCAACACAGCGCTAAAGTGTCTCAGCTCGGCCAAA 2450
Qy 812 AsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCys 831
Db 2451 AATAGAGTGGGAGTAAAGACATGATATCCCTTGGGAATGCTTGAAGTAGTCAATGTGT 2510
Qy 832 AlaLeuGlnLysIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeu 851
Db 2511 CTACTCAAAAGTTGATCTGGAACACTGTGCGCTCACACCTGCGCAATGTCTCTCCAG 2570
Qy 852 AlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeu 871
Db 2571 GTCTCAGCCCTTTTCAGCAACCAAGACTTGACACACTGTGCTGTCTGCAAAACACGCTG 2630
Qy 872 GlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGln 891
Db 2631 GGGACTGAAGGAGTGCAACAGCTGTCTGAGTCTCTGAGGAATCCAGAAATGTCTCTCCAG 2690
Qy 892 ArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAla 911
Db 2691 CGCTGATGATCAATCACTGCAACATTTAGATGATGCTTATGGCTTCTCGCAATGAGA 2750
Qy 912 LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsn 931
Db 2751 CTTGCAAAACACAAAGCTGACCCACTGAGCTGACCTGAAACCCCTAGGGATGGT 2810
Qy 932 GlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGTACTGTGTGAAGCTTTAAAGGAACCTTACTTGTACCTTCAAGAACTGGAA 2870

```
QY 952 LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db 2871 CTATGGGACTGCACACTCACAGAACTGCTCGGAGGACCTGGCCTGTATGATCAACA 2930
QY 972 SerArgHisLeuSerLeuAspLeuThrAspAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAGACCTTAAAGATTGATCTTGGTAAACACGCCCTGGGTGACAAAGGAGTCATA 2990
QY 992 AlaLeuCysGluGlyLeuGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGACTGAAGCAAGTAGCAGCTCCCTGAGGAGACTTGGGTGGGGCA 3050
QY 1012 CysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAATTGACTTCCAAATTGCTGAGGCAATTGTCAATTGGCCATCTCTTGAACCTCAC 3110
QY 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCys 1051
Db 3111 CTGAACAGCCTTAACCTGTGAGGAAATGACTTCAGTACATCGGGATGTTGAAGCTGTGC 3170
QY 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrp 1071
Db 3171 TCTGGTTTCCAATGCCCTGTCTTAACCTGGGGATAATTGGCTGTGGAAACAGAGTAC 3230
QY 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArgValValle 1091
Db 3231 TATGCCCGAGTGAGAAGACAGCTGGAGGAAGTTGAGTTTGTCAAGCCCAAGCTGCTGATT 3290
QY 1092 AspGlySerTrpHisSerPheAsnGluAspArgHis 1104
Db 3291 GATGGTGAATTGGTATGTAGTGATGATGAAGATGACCCGAAC 3329
RESULT 12
AAD49001
ID AAD49001 standard; cDNA; 3447 BP.
XX AAD49001;
XX
XX 07-MAR-2003 (first entry)
XX
XX Mouse MATER cDNA.
XX
XX Mouse; MATER protein; infertility; fertility; contraceptive agent;
XX gene therapy; gene; ss.
XX
XX Mus sp.
XX
XX Key
XX CDS
XX
XX Location/Qualifiers
XX 9..3344
XX /tag= a
XX /product= "Mouse MATER protein"
XX /transl_except= (pos:648..650, aa:Thr)
XX /transl_except= (pos:714..716, aa:Thr)
XX /transl_except= (pos:768..770, aa:Thr)
XX /transl_except= (pos:882..884, aa:Thr)
XX /transl_except= (pos:1359..1361, aa:Thr)
XX /transl_except= (pos:1554..1556, aa:Thr)
XX /transl_except= (pos:1740..1742, aa:Thr)
XX /transl_except= (pos:1869..1871, aa:Thr)
XX /transl_except= (pos:2046..2048, aa:Thr)
XX /transl_except= (pos:2049..2051, aa:Thr)
XX /transl_except= (pos:2229..2231, aa:Thr)
XX /transl_except= (pos:3216..3218, aa:Thr)
XX /transl_except= (pos:3300..3302, aa:Thr)
XX /transl_except= (pos:3330..3332, aa:Thr)
XX /transl_except= (pos:3333..3335, aa:Thr)
XX /note= "No start and stop codon"
XX /partial
XX
XX W0200281492-A1.
XX
XX 17-OCT-2002.
XX
```

```
PF 29-MAR-2002; 2002WO-US0009776.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2003-058494/05.
XX
XX P-PSDB; AAE31748.
XX
XX New isolated variant MATER proteins and nucleic acids, useful for
XX diagnosing, prognosing and treating infertility and reduced fertility,
XX and as contraceptive agents.
XX
XX Example 22; Page 83-85; 110pp; English.
XX
XX The present invention relates to novel MATER proteins and polynucleotides
XX encoding such proteins. The MATER proteins are essential to fertility.
XX Sequences of the invention are useful for diagnosing, prognosing and
XX treating infertility, reduced fertility and as contraceptive agents. They
XX are also useful in gene therapy. The method is useful for detecting a
XX predisposition to or pre-symptomatic screening of an individual for
XX infertility or reduced fertility. The present sequence is mouse MATER
XX cDNA
XX
XX Sequence 3447 BP; 1000 A; 783 C; 843 G; 821 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2,94e-204 Length: 3447
XX Score: 2688.00 Matches: 554
XX Percent Similarity: 66.49% Conservative: 186
XX Best Local Similarity: 49.78% Mismatches: 293
XX Query Match: 38.00% Indels: 81
XX DB: 7 Gaps: 11
XX
XX US-10-066-521-6 (1-1344) x AAD49001 (1-3447)
XX
XX QY 15 GlnTrpCysLeuTyrgLysLeuAspLysGluGluPheGlnThrPhelysGluLeuLys 34
XX Db 163 CAATGGTCTCTCCA-GAAAAGACAGTAAAGCAATCTTGAAGCAGCCTGATGGAAGAG 221
XX
XX QY 35 LysLysSerSerGluSerThrCysSerSileProGlnPheGluIleGluAsnAlaAsn 54
XX Db 222 GAACAGAGAGTCAGAAAGCACA----- 242
XX
XX QY 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrgLysAlaSerLeuAlaTrpAlaThr 74
XX Db 242 ----- 242
XX
XX QY 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
XX Db 243 ---ATGCTCTCTTCAGAAATGTCAGTAGAGCAATCTCTGAAGACAGCTGATCAGAAGAA 299
XX
XX QY 95 MetLysLysIleSerGln----- 100
XX Db 300 GTGGAAACAGCGCTCAGAAAGAAAATGACTCTCCAGAAAAACGACGATAAATCAATCCAG 359
XX
XX QY 101 -----AlaMetGluGlnGluGluAlaThrAlaAlaGluThrGluGluGln 115
XX Db 360 AAAGACCAACGAGCAGAGCAGGACATCAGAAACCTTACAACTTAAGGAGGAAGAGT 419
XX
XX QY 116 GluIleSerGlnAlaMetGluGlnGluGluAlaThrAlaAlaGluThrGluGluGlnGly 135
XX Db 420 GAAGTCACAGAGCA-----GATAAGAT 443
XX
XX QY 136 HisGlyGlyAspThrTrpAspTyrgLysSerHisValMetThrLysPheAlaGluGluGlu 155
XX Db 444 AATGGAGGTGACTTACAAGACTACAGGCCCATGTGATTGCTAAGTTTCGACACACAAGGTG 503
XX
XX QY 156 AspValArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
XX Db 504 GATCTACAC-----TATGACAGCCCGCAGAGATGAAATTTGTTCT 542
```



```
Db 2691 CGCGTGATGTAACACTGACGCAACATTTAGATGATGCTTATGGCTTCTCGCAATGAGA 2750
Qy 912 LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuMetAsnProValGluAspAsn 931
Db 2751 CTTCGAACACACAAAGCTGACCCACTGACCTGACCATGACATGAAACCCGCTAGGGGATGGT 2810
Qy 932 GlyValLysLeuLeuGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGCTACTGTGTGAAGCTTTAAAGGAACCTACTTGTACCTTCAAGAACTGGAA 2870
Qy 952 LeuValLysCysHisLeuThrAlaLysCysGluSerLeuSerCysValLysSerArg 971
Db 2871 CTAGTGGACTGCAACTCACAGAACTGCTGCGAGGACCTGGCTGTATGATCACACA 2930
Qy 972 SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAGCACTTAAAGTTTGGATCTTGTAAACAGCCCTGGGTGACAAAGGATGATATA 2990
Qy 992 AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGACTGAAGCAAAAGTAGCAGCTCCCTGAGGAGACTTGGGTGGGGCA 3050
Qy 1012 CysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAGTTGACTTCCAATTGCTGTGAGGCATTGTCAATTGGCCATCTCTTGAACCCCTCAC 3110
Qy 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetLysLeuCys 1051
Db 3111 CTGAACAGCTTAAACCTGTGTAAGTAATGACTTCAGTACATCGGGATGTTGAAGCTGTGC 3170
Qy 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyr 1071
Db 3171 TCTGCTTCCAATGCCCTGTCTCTAACTGGGGATAATTTGGCTGTGGAAGCAGGAGTAC 3230
Qy 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIle 1091
Db 3231 TATGCCCGAGTGAGGAAGACAGCTGAGGAAGTTGAGTTTGTCAAGCCCAAGCTGTGTATT 3290
Qy 1092 AspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3291 GATGTGATTGTGTATGCTAGTGTGAAGATGACCGAAAC 3329

RESULT 13
ABK48609
ID ABK48609 standard; cDNA; 1157 BP.
XX AC ABK48609;
XX DT 13-AUG-2002 (first entry)
XX DE Human MATER cDNA fragment #1.
XX KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; WATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX OS Homo sapiens.
XX FH Key
XX CDS Location/Qualifiers
FT 1..1155
FT /*tag= a
FT /product= "Human MATER"
FT /partial
FT /note= "No start or stop codon shown"
XX WW020232955-A1.
XX PN
XX PD 25-APR-2002.
XX PF 04-APR-2001; 2001WO-US010981.
XX PP
```

```
PR 18-OCT-2000; 2000US-0241510P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Nelson LM, Tong Z;
XX WPI; 2002-454595/48.
XX P-PSDB; AAU79523.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
XX fertility, or as a contraceptive.
XX
XX Claim 11; Page 68-69; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy Maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA
CC fragment #1, which was isolated from a human cDNA library
XX
SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,16e-150 Length: 1157
Score: 2005.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.34% Indels: 0
DB: Gaps: 0

US-10-066-521-6 (1-1344) x ABK48609 (1-1157)
Qy 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuPheIleIleAspGly 271
Db 1 CAAGCTCCGGTGACGGAGATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGACGGT 60
Qy 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291
Db 61 TTCATGACCTGGGCTCTGTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG 120
Qy 292 LysGlnProProPheThrLeuIleArgSerLeuArgLysValLeuLeuProGluSer 311
Db 121 AAGCAGCTCCGTTTCACTCATACCGCAGTCTGTCTGAGGAAGGCTCTCTCTCCCTGAGTCC 180
Qy 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
Db 181 TTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTGTCT 240
Qy 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu 351
Db 241 CCCGTTACCTGTGTAGTGGAGTAATCTCCGGGACCAAGATCCACTTGTCTCTTGAG 300
Qy 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
Db 301 CGCGGGATTGGTGAGCATCAGACACACAAAGGGTTGCGTGCATCATGAACACCGTGAG 360
Qy 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACCATGTCAGGTGCCCGCGGTGGGCTCTCTCATCTCGTGGCCCTTCAGCTG 420
```

```
QY 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
DB 421 CAGGACGTGGTGGGAGAGGTGCCCCCTTCAACCAAGCGCTCACAGGCTCGACGCC 480
QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431
DB 481 GCTTTTGTGTTTCATCAGCTCACCCCTCAGAGCGTGGTCCGGCGCTGTCTCAATCTGGAG 540
QY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
DB 541 GAAAGAGTTGTCTGTAAGCGCTTCTGCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAG 600
QY 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
DB 601 TCAGTGTGTGAGCGGTGACGACTCATGGTTCAAGACTCGGGGAGTCTGAGCTCCGTGCT 660
QY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyTrpThrPhePhe 491
DB 661 CTGTTTTCATGACATCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTC 720
QY 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyTrpValLeuGluGlyLeuGluIle 511
DB 721 CACCTCAGTCTCCAGGACTTCTGTGCGCCTTGTACTACGTGTAGAGGCGCTGGAATC 780
QY 512 GluProAlaLeuCysProLeuTyTrpValGluLysThrLysArgSerMetGluLeuLysGln 531
DB 781 GAGCCAGCTCTCTGCGCTCTGTAGCTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAG 840
QY 532 AlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGlu 551
DB 841 GCAGGCTTCCATATCCACTCGCTTGGATGAAGCGTTTCTGTTGGCCTCGTGAGCGAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
DB 901 GACGTAAAGAGGCCACTGGAGGTCTCTGGGCTGTCCCGTTCCTCCCTGGGGTGAAGCAG 960
QY 572 LysLeuLeuHisTrpValSerLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
DB 961 AAGCTTCTGCACCTGGGTCTCTCTGTGGGTGAGAGCTTAATGCCACCCAGGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
DB 1021 ACCCTGGAGCGCTTCCACTGCTTTTTCGAGACTCAGACAAAGAGTTTGTTCGCTGGCA 1080
QY 612 LeuAsnSerPheGlnGluValTrpProLysGlnGlnAsnGlnLeuAspLeuIleAlaSer 631
DB 1081 TTAACAGCTTCCAAAGAAGTGGCTTCCGATTAAACCAAGAACCTGGACTTGATAGCATCT 1140
QY 632 SerPheCysLeuGln 636
DB 1141 TCCTTCTGCGCTCCAG 1155
RESULT 14
AAD48999
ID AAD48999 standard; cDNA; 1157 BP.
XX
AC AAD48999;
XX
DT 07-MAR-2003 (first entry)
XX
DE Human MATER cDNA fragment 1.
XX
KW Human; MATER protein; infertility; fertility; contraceptive agent;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1155
FT /tag= a
FT /product= "Human MATER protein fragment 1"
FT /transl_except= (pos:112..114, aa:Thr)
```

```
FT /transl_except= (pos:589..591, aa:Thr)
FT /transl_except= (pos:865..867, aa:Thr)
FT /transl_except= (pos:973..975, aa:Thr)
FT /transl_except= (pos:1102..1104, aa:Thr)
FT /note= "No start and stop codon"
FT /partial
XX
PN WO200281492-A1.
XX
PD 17-OCT-2002.
XX
PF 29-MAR-2002; 2002WO-US009776.
XX
PR 04-APR-2001; 2001WO-US010981.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
WP1; 2003-058494/05.
P-PSDB; AAE31746.
XX
New isolated variant MATER proteins and nucleic acids, useful for
diagnosing, prognosing and treating infertility and reduced fertility,
and as contraceptive agents.
XX
Example 2; Page 79-80; 110pp; English.
XX
The present invention relates to novel MATER proteins and polynucleotides
encoding such proteins. The MATER proteins are essential to fertility.
Sequences of the invention are useful for diagnosing, prognosing and
treating infertility, reduced fertility and as contraceptive agents. They
are also useful in gene therapy. The method is useful for detecting a
CC predisposition to or pre-symptomatic screening of an individual for
infertility or reduced fertility. The present sequence is human MATER
cDNA fragment 1
XX
SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	2,16e-150	Length:	1157
Score:	2005.00	Matches:	385
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	28.34%	Indels:	0
DB:	7	Gaps:	0

US-10-066-521-6 (1-1344) x AAD48999 (1-1157)

QY	252	GlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGly	271
DB	1	CAAGCTCCCGTGCAGGAGATCATGTCGCCAGCAGAAAGGCTGTGTTCATCATTTGACGGT	60
QY	272	PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu	291
DB	61	TTGATGACCTGGCTCTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG	120
QY	292	LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer	311
DB	121	AAGCAGCCTCCGTTCAACCTCATACGAGTCTGCTGAGGAAGTCTGCTCCCTGAGTCC	180
QY	312	PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer	331
DB	181	TTCTCTGATCGTCAACGTCAGAGACGCTGGGCAACAGAAAGCTCAAGTCAGAGGTCGT	240
QY	332	ProArgTyTrpLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu	351
DB	241	CCCCGTTACCTGTGTAGTAGGAATCTCCGGGGACAAAGAAATCCACTTGTCTCTTGAG	300
QY	352	ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu	371
DB	301	CGCGGGATTGGTGAGCATCAGAAAGACACAGAGGTTGGTGGCATCATGAACACCTGAG	360


```
QY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACGATGCGAGTGCCTGCGCTGCTCTCATCTGCGTGCCTCGAGCTG 420
QY 392 GlnAspValValGlySerValAlaProPheGlnGlnThrLeuThrGlyLeuHisAla 411
Db 421 CAGGACGTGGTGGGAGAGGCTGCGCCCTTCAACCAAGCGCTACAGGCTGCGAGCC 480
QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValAlaArgCysLeuAsnLeuGlu 431
Db 481 GCTTTTGTGTTTCATCAGCTCAGCCCTCAGGCGGTGGTCCGCGCTGCTCAATCTGGAG 540
QY 432 GluArgValValLeuLeuArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
Db 541 GAAAGAGTGTCTCGAGCGCTTCTGCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAG 600
QY 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
Db 601 TCAGTGTGTGACGGTGAGCCTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCT 660
QY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePhe 491
Db 661 CTGTTTCATGACATCCTCTCCACAGACCACTGTGAGGAGTACTACACTTCTTC 720
QY 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIle 511
Db 721 CACCTCAGTCTCCAGGACTTCTGTGCGCTTGTACTAGTGTGTAGAGGCTCGAAATC 780
QY 512 GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCAGAGCTCTGCGCTCTGTACGTTGAGAGAGCAAAAGAGGTCATGGAGCTTAAACAG 840
QY 532 AlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGGCTTCCATATCATCTGCTTGGATGAGAGCGTTTCTTGTGGCTCTGAGCGAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GACGTAAAGAGGCGCACTGGAGTCTCTGCTGGCTGCTCCGTTCCCTCGGGGTGAAGCAG 960
QY 572 LysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGCATGGTCTCTCTGTGGTGCAGAGCCCTAATGCCACCCACCGAGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGGAGCGCTTCCATCTCTTTCGAGACTCAAGACAAAGAGTTGTCGCTTGCCA 1080
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1081 TTAAACAGCTTCCAAGAAGTGTGGCTTCCGATTAAACAGAACCTGGACTTGATAGCATCT 1140
QY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 15
AAL4366
ID AAL44366 standard; cDNA; 3368 BP.
AC AAL44366;
XX
DT 31-OCT-2002 (first entry)
XX
DE Human PYRIN-3 cDNA sequence.
KW Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-3.
```

```
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 70..3054
FT /*tag= a
FT /product= "Human PYRIN-3 protein"
XX
XX WO200261049-A2.
XX
XX 08-AUG-2002.
XX
XX 31-JAN-2002; 2002WO-US002967.
XX
XX 31-JAN-2001; 2001US-0265231P.
XX
XX 10-SEP-2001; 2001US-0318645P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX (AMHP ) WYETH.
XX
XX Bertin J, Wang W, Blatcher M;
XX
XX WPI; 2002-627477/67.
XX
XX P-PSDB; AAO15593.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
XX diagnosing stress-related, apoptotic and inflammatory responses, or for
XX treating inflammatory and immune system disorders, cancers, or
XX neurological diseases.
XX
XX Claim 4; Fig 11; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
XX useful for modulating and diagnosing stress-related, apoptotic and
XX inflammatory responses. The PYRIN protein and DNA sequences are useful
XX for treating: inflammatory disorders and immune system disorders (e.g.
XX Crohn's disease, reactive arthritis, multiple sclerosis, contact
XX dermatitis, psoriasis, graft rejection, allergies, viral infections and
XX bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
XX (e.g. systemic lupus erythematosus and arthritis); and neurological
XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
XX protein and DNA sequences may also be used in screening assays, detection
XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
XX predictive medicine (e.g. diagnostic assays, clinical trials and
XX pharmacogenomics) and transcription profiling. The present DNA sequence
XX encodes the human PYRIN-3 protein
XX
XX SQ Sequence 3368 BP; 900 A; 803 C; 828 G; 837 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.62e-104 Length: 3368
Score: 1435.00 Matches: 349
Percent Similarity: 48.05% Conservative: 204
Best Local Similarity: 30.32% Mismatches: 448
Query Match: 20.29% Indels: 150
DB: 6 Gaps: 15

US-10-066-521-6 (1-1344) x AAL44366 (1-3368)
QY 9 PheSerSerTyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThr 28
Db 85 TTCTCTGATTTGGTCTTATGTGGTATCTGGAGGAGCTCAAAAGGAGGAGTTCAGGAA 144
QY 29 PheLysGluLeuLeuLysLysSerSerSerThrThrCysSerIleProGlnPhe 48
Db 145 TTTAAAGACATCTCAAGCAAAATGACTTTCAGCTTCAACTCAAGCAGATTCCTGGACT 204
QY 49 GluIleGluAsnAlaAsnValGluCysLeuAlaLeuLeuHisGluTyrTyrGlyAla 68
Db 205 GAGGTCAAAAAGAGCATCCCGGGAAGAACTTGCACCAACCTTGTATCAAGCTACTGAAGAA 264
QY 69 SerLeuAlaTrpAlaThrSerIlePheGluAsnMetAsnLeuAlaArgThrLeuSer 88
```

Db 265 CAACAAAGCTTGGAACATAACCTTAAGAATCTTTCAAAAGATG----- 306
Qy 89 GluLysAlaArgAspMetLysLysLysSerGlnAlaMetGluGlnGluGlyAlaThr 108
Db 307 -----GATAGAAAGGATCTCTGCATGAAGGTCATGAGGGAGAGAACAGGA 351
Qy 109 AlaAlaGluThrGluGlnGluLysSerGlnAlaMetGluGlnGluGlyAlaThrAla 128
Db 352 TACACAAAGACCTATCAGCTCAGCAAGCAGCAAGCAGAAATTCACCGCG----- 396
Qy 129 AlaGluThrGluGlnGluGlnGlyHisGlyGlyAspThrTyrAspTyrLysSerHisValMet 148
Db 397 -----TTATGGTCCAGCAAGTCTGTCACTGAG 423
Qy 149 ThrLysPheAlaGluGluAspValArgArgSerPheGluAsnThrAlaAlaAspTyr 168
Db 424 ATTCACCTATATCTTTGAGGAGGAGTCAAGCAAGAA----- 459
Qy 169 ProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTyrGlyPheArgProArg 188
Db 460 ---GAATGTGACCATTTGGACCGCTTTTGTCTCCCAAGGAAACTGGGAAACAGCCACGT 516
Qy 189 ThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleVal 208
Db 517 ACAGTGATTATTCAAGGACCAACAGGAATTCGAAACACGACACTCTGTGATGAAGCTGATG 576
Qy 209 LeuCysTyrAlaGlnGlyLysLeuTyrGlnGlyMetPheSerTyrValPhePheLeuPro 228
Db 577 ATGGCTCGGTCCGGAACAAAGATCTTTGGGATAGTTCTCTGTACAGGTCTTATTTCTGCG 636
Qy 229 ValArgGluMetGlnArgLysLysSerValThrGluPheIleSerArgGluTyr 248
Db 637 TGCAGAGAACTGAGGGAGTTCGCCCAACGAGTTTGGCTGACTTGTATTTCCAGAGATGG 696
Qy 249 ProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIle 268
Db 697 CTGACCCCGCTGCTCTATTAACAGAGATCTGTGTCTCAACCGGAGAGACTCTTGTTCGTC 756
Qy 269 IleAspGlyPheAspAspLeuGlySerValLeuAsnAsn---AspThrLysLeuCysLys 287
Db 757 ATCGACAGCTTCGAAGAGCTCGACGGCGGCTGTGAACGAACCGGATTCGGATCTGTGTGGT 816
Qy 288 AspTyrAlaGluLysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeu 307
Db 817 GACTTGATGGAACACGGCCGGTCAGGTGCTCTGAGCAGATTTCCTGAGGAAGAAGATG 876
Qy 308 LeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSer 327
Db 877 CTCCCGGAGGCTCCCTGCTCATCGCATCAACCCGCTGTGCCGGAAGGAGCTCCGGGAT 936
Qy 328 GluValValSerProArgTyrLeuLeuValArgGlyLysSerGlyGluGlnArgIleHis 347
Db 937 CAGGTGACGATCTCAGAAATCTACACGCCCGGGGATTCACACGAGAGTGATAGTTAGTG 996
Qy 348 LeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMet 367
Db 997 TATTTCTGCTTTCTTCAAGACCCGGAAGACCCATGGAAGCCTTCAATCTTGTGAAGA 1056
Qy 368 AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal 387
Db 1057 GAAAGTGAACAGCTGTTTTCATATATGCCAAATCCCGCTCTCTGCTGATCTGTGTACC 1116
Qy 388 AlaLeuGlnLeuGlnAspValValGlyLysSerValAlaProPheAsnGlnThrLeuThr 407
Db 1117 AGTCTGAAGCAAGATGTCGAAAGGAAAGAACCTGCGCCCTGCACCTGCCAGAGACTTACC 1176
Qy 408 GlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCys 427
Db 1177 TCTGTGTACTCTCTTTCGTTCTTATCTGTTTCACTGAGGTGCGGAGGCCCGCACT 1236
Qy 428 LeuAsnLeuGluArgValValLeuLysArgPheCysArgMetAlaValGluGlyVal 447
Db 447 ----- 447

Db 1237 CCGCAAAACCCAGCACACAG-----CTGAAGGCCCTGTGCTCCCTGCTGCAGAGGGTATG 1290
Qy 448 TrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSer 467
Db 1291 TGGACAGACACATTTGAGTTTGTGAAGACGACCTCGGAGAAATGGGTTGTGACGT 1350
Qy 468 GluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGlyGluTyr 487
Db 1351 GACATCCCTCGGTGTGGCCACCAAGATACTTCTGAAGTACGGGAGCGTGAAGCTCC 1410
Qy 488 TyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGlu 507
Db 1411 TAGGTGTTCTCCACGTGTATCCAGGAGTTCTGTGCGCCCTGTATTCTATTGCTCAAG 1470
Qy 508 GlyLeuGluIleGluPro-----AlaLeuCysProLeuTyr 519
Db 1471 AGCCACCTTGATCATCTCAACCAGCTGTGAGATGTGTACAGGAATTGTAGTTGCCAAT 1530
Qy 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisLeuSerLeu 539
Db 1531 TTTGAAAACCAAGAGAGCA-----CATTGAT 1560
Qy 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
Db 1561 TTTTGGGGTGTCTTAAGTGGCTTTTAAATAAAAGCAACAGAAAACTGGATGCG 1620
Qy 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTyrValSerLeu 579
Db 1621 TTTTGGGTTCCAACTGTCCCAAGAGATAAAGACAGCAAAATTCACAGTGCCTGAAGAGC 1680
Qy 580 LeuGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 1681 TTAGGAGCGGTGGCAATCTCAGGACAGGTGGATCTCTGGCGATATTTACTGTCTC 1740
Qy 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTyr 619
Db 1741 TTTGAAATGCAGGATCTGCTCTTTGTGAAGCAGCGCAGTGAACCTCTCCAAGAAGCTAAC 1800
Qy 620 LeuProIleAsnGlnAsnLeuAspIleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 1801 TTTTCATATTATTGACAACGTGACTTGTGGTGTCTGCTACTGTCTTAAATACTGCTCC 1860
Qy 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
Db 1861 AGCTTGAGGAACTCTGTTTCCGTTCAAAATGTCTTTAAGAAAGAGATGAACACAGC 1920
Qy 660 AlaCysProValValProLeuTyrMetArgAspLysThrLeuIleGluGlnTyrGlu 679
Db 1921 TCT-----ACGTCCGATTACAGCCTCATC-----TGTGGCAT 1953
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 1954 CACATCTGCTCTGTGTCACCAAGCGGCGACCTCAGAGAGCTCCAGGTGCGAGGACAGC 2013
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2014 ACCTCAGCGAGTCCAGCTTTGTGACCTGGTTAACCTGAGTCATCCAGCAACTCTCTCGT 2073
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTyrArg 739
Db 2074 CTTTCAAGCTTTGGAATAAATACGGTTCTTTCTGGCCAGAGTGTCTGCTCTTTGAG 2133
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 2134 GTGCTCTTTTATCAGCCAGACTTGAATACCTGAGCTTCACCCCTCACGAAACTCTCTCGT 2193
Qy 760 GluAspValArgMetAlaCysGluAlaLysHisProLysCysLeuLeuGluSerLeu 779
Db 2194 GATGATCAGGTCCCTCTGTGATGCTTGAACACTACCCA----- 2232
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2232 ----- 2232

Search completed: July 16, 2004, 00:59:44
Job time : 1282 secs

```
QY 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 2233 -----GCAGCAAC----- 2241
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 2242 -----GTCAAAGAGCTAGCGCTGGTA 2262
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2263 AATGTGACCTCTACCATGATTGTGAATGCTTGTGCGCTTCTAACCAACAACAAG 2322
QY 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2323 AAGCTGACGTATCTGAATGATCTCTGCAACACAGTTA---GACACAGGCGTCCCTTTTG 2379
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2380 TGTGAAGCCCTGTGCAGCCCAACACACACGCTCTGCTATACCTGATGTTGGCTTTCTGCCAC 2439
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2440 CTCAGCGAGAGTGTGCAATACATCTCTGAATGCTTCTGCGTAAACAGAGCGTGGC 2499
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLysGluVal 939
Db 2500 TATCTAGACCTCAGTCCCAATGCTCTGAAGACGACGAGGACTGMAAACTCTCTGCGAGGCC 2559
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 2560 TTGAACATCCGGACTGCTGCTGGATTCACCTGTGTTGGTAAATGTTTATCCTGCT 2619
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 2620 GCTGCTGTGAAGACCTCGCTCTGCTCATCAGCAATCAAAACCTGAAGATTCTGCAA 2679
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 2680 ATTGGGTGCAATGAATCGAGATGTGGTGTGCAGCTGTTGTGCGGGCTCTGACGCAT 2739
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 2740 ACGGATTCCCGCTAGAGATTCTGGGTTGAAGAAATGCGGTTAAACGAGCACCTGCTGT 2799
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 2800 AAGGATCTCGCGTCTGTTCTCACCTGCAGTAAGACCCCTGCAGCAGCTCAACCTGACCTTG 2859
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 2860 AACACCTTGGACCACACAGGGGTGGTGTACTCTGTGAGGCGCTGAGACACCCAGAGTGT 2919
QY 1060 AsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 2920 GCCCTGCAGGTGCTCGGCTCGAGAAAACATGATTTGATGAGAAAACCCAGGCACCTTCTG 2979
QY 1080 GluGluValGlnLeuLysProArgValIleLeuAsp----- 1092
Db 2980 ACGGCTGAGGAAGAGAGAAATCCTTAACCTGACCATGACAGATGCTGTGACACCAATCACA 3039
QY 1093 -----GlySerTrpHisSerPheAspGluAspArgHisLys 1105
Db 3040 AGGGTAGAGATCTGATTCGGAGGAACCTGG----- 3069
QY 1106 IleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCysAlaLeuLeuTrpGly 1125
Db 3070 ---GCTCTGACTCGAACAACCTGCAAGGACAGGAGCTGGAC---CGTTACTTACATGAC 3123
QY 1126 MetAsnProGluGlnLysLysArgValSerLeu 1136
Db 3124 ACTGCACCCAGGATACAAATCATTTGACATC 3156
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 20:38:28 ; Search time 200 Seconds
(without alignments)
3729.270 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074

Sequence: 1 MEGDKSLTFSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10066521/runat_13072004.122213.1111/app_query.fasta_1.1543
-DB=Issued Patents_NA -QFWT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521 @CGN 1 1 107 @runat_13072004.122213.1111 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=1120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699.5	9.9	1566	US-09-833-381-1975	Sequence 1975, Ap
2	630.5	8.9	1374	US-08-910-731-3	Sequence 3, Appli
3	630.5	8.9	1374	US-08-795-395-3	Sequence 3, Appli
4	621	8.8	1371	US-08-910-731-1	Sequence 1, Appli
5	621	8.8	1371	US-08-795-395-1	Sequence 1, Appli
6	612	8.7	1371	US-08-910-731-7	Sequence 7, Appli
7	607.5	8.6	1386	US-08-910-731-5	Sequence 5, Appli
8	458	6.5	1515	US-09-833-381-1904	Sequence 1904, Ap
9	458	6.5	1515	US-09-833-381-1905	Sequence 1905, Ap
10	394.5	5.6	1650	US-09-833-381-1976	Sequence 1976, Ap
11	379.5	5.4	4141	US-09-245-281-42	Sequence 42, Appl
12	379.5	5.4	4141	US-09-207-359B-42	Sequence 42, Appl

13	379.5	5.4	4141	4	US-09-340-620A-42	Sequence 42, Appl
14	379.5	5.4	4141	4	US-09-865-364-42	Sequence 42, Appl
15	359.5	5.1	3382	4	US-09-099-041A-7	Sequence 7, Appli
16	359.5	5.1	3382	4	US-09-245-281-7	Sequence 7, Appli
17	359.5	5.1	3382	4	US-09-207-359B-7	Sequence 7, Appli
18	359.5	5.1	3382	4	US-09-340-620A-7	Sequence 7, Appli
19	359.5	5.1	3382	4	US-09-865-364-7	Sequence 7, Appli
20	347	4.9	2859	4	US-09-099-041A-9	Sequence 9, Appli
21	347	4.9	2859	4	US-09-245-281-9	Sequence 9, Appli
22	347	4.9	2859	4	US-09-207-359B-9	Sequence 9, Appli
23	347	4.9	2859	4	US-09-340-620A-9	Sequence 9, Appli
24	347	4.9	2859	4	US-09-865-364-9	Sequence 9, Appli
25	327	4.6	4543	2	US-08-519-547A-5	Sequence 5, Appli
26	319.5	4.5	4346	4	US-09-064-199-12	Sequence 12, Appl
27	319.5	4.5	4366	4	US-09-064-199-14	Sequence 14, Appl
28	319.5	4.5	4418	4	US-09-064-199-13	Sequence 13, Appl
29	319.5	4.5	4431	4	US-09-064-199-8	Sequence 8, Appli
30	319.5	4.5	4441	4	US-09-641-999-2	Sequence 2, Appli
31	319.5	4.5	4441	4	US-09-064-199-10	Sequence 10, Appl
32	319.5	4.5	4549	4	US-09-064-199-9	Sequence 9, Appli
33	319.5	4.5	4564	4	US-09-064-199-2	Sequence 2, Appli
34	319.5	4.5	4649	4	US-09-064-199-11	Sequence 11, Appl
35	319.5	4.5	4746	4	US-09-064-199-7	Sequence 7, Appli
36	319.5	4.5	5105	4	US-09-064-199-3	Sequence 3, Appli
37	319.5	4.5	5463	4	US-09-064-199-1	Sequence 1, Appli
38	309	4.4	3393	1	US-08-295-502-1	Sequence 1, Appli
39	309	4.4	3393	5	PCT-US95-10691-1	Sequence 1, Appli
40	307	4.3	4302	4	US-09-245-281-38	Sequence 38, Appl
41	307	4.3	4302	4	US-09-207-359B-38	Sequence 38, Appl
42	307	4.3	4302	4	US-09-340-620A-38	Sequence 38, Appl
43	307	4.3	4302	4	US-09-865-364-38	Sequence 38, Appl
44	296.5	4.2	2480	4	US-09-064-199-15	Sequence 15, Appl
45	235	3.3	3116	4	US-09-904-615-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-833-381-1975
; Sequence 1975, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1975
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1975

Alignment Scores:
Pred. No.: 1.51e-63 Length: 1566
Score: 699.50 Matches: 178
Percent Similarity: 51.56% Conservative: 103
Best Local Similarity: 32.66% Mismatches: 201
Query Match: 9.89% Indels: 63
DB: 4 Gaps: 15

US-10-066-521-6 (1-1344) x US-09-833-381-1975 (1-1566)

Oy	139	AspThrTrpAspTyr-----LysSerHisValMetThrLysPheAlaGluGlu 154
Db	19	GAGACATGGGACACACATCAGTGGCGCTTAAGACCCATATATATCCGT----- 66
Oy	155	GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrIleu 174


```

757 CTGAGACTCTGTGGCTCTGGAGACTGTGATGTCACTGCAGAAGCGCTGCCAAGGACCTGTGC 816
910 LeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGlu 929
    ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
817 CGTGTCTCAGAGCCAAAGCAGACGAGCCTGAAGGAACCTCAGCCTAGCTGGCAANTGAGTCGAAG 876
930 AspAsnGlyValIysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAsp 949
    ||| ||| :::||||| ||| ||| :::||||| ||| ||| ||||| ||| ||||| ||| |||||
877 GATGAGGGTGCCCAACTGCTGTGTGAGAGCCTGTATTAGAGCCTGGCTGTGACGTGGAGTCA 936
950 LeuGluLeuValIysCysHisLeuThrAlaAlaCysGluSerLeuSerCysValIle 969
    ||| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
937 CTGTGGGTAAAGACCTGTAGCCTCACAGCTGCTCTTTGTCCTCCCACCTTCTGCTCGGTGTG 996
970 SerArgSerArgHisLeuIysSerLeuAspLeuThrAspAlaLeuGlyAspGlyGly 989
    ||| ::||:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
997 ACCAAAACAAGTTCTCTGTTTGAGTTTGCAAATGAGCAGCAACCGCTGGGAGACTCGGGA 1056
990 ValAlaAlaLeuCysGluGlyLeuLysGlnIysAsnSerValLeuThrArgLeuGlyLeu 1009
    ||| ||||| ||||| ||||| ||||| ::||:: ||||| ||||| ||||| ||||| ||||| |||||
1057 GTCGTGGAGCTTTGCAAGGCCCTTGGGCTATCCCGACACAGTGTGCTGTGCTTTGGCTG 1116
1010 LysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsn 1029
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1117 GGAGACTGTGATGTGACAGACAGTGGCTGCAGACGCTTCCACTGTCTGCTGGCCCAAC 1176
1030 ArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProIysGlyMetMetLys 1049
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1177 CGCAGCTTGAGGGAACCTGGACCTCAGTAACAACCTGCATGGGGGACACAGGTGTCTACAA 1236
1050 LeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeu----- 1066
    ||| ::| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1237 CTGCTGGAGAGCCTCAACACAGCCAGCTGCATCCTTCAGCAGCTGTGTCTGTATGACATT 1296
1067 ---TrpIysrPrGlnTyPrValGlnIleArgIysLeuGlu 1081
    ||||| ||||| ::||:: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1297 TACTGGACGATGAGGTGGGAAGACCAGCTTCGGGCCCTGGAGGAGGAA 1344

RESULT 3
US-08-795-395-3
; Sequence 3, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600

```


; NAME: ESMOND, ROBERT W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0942.3440003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1371 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1368
 ; US-08-910-731-1

Alignment Scores:

Pred. No.: 2,44e-55 Length: 1371
 Score: 621.00 Matches: 151
 Percent Similarity: 50.11% Conservative: 86
 Best Local Similarity: 31.92% Mismatches: 210
 Query Match: 8.78% Indels: 26
 DB: 2 Gaps: 5

US-10-066-521-6 (1-1344) x US-08-910-731-1 (1-1371)

```

Qy 625 AsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIle 644
Db 4 AACCTCGATATT-----CATTCGAGAGCTGAGCGAGCC 39
Qy 645 ArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValVal 664
Db 40 CGG-----TGACAGAGCTCTCCCTCGCTC 66
Qy 665 ProLeuTrp-----MetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 67 CAGCAGATATGAGTGGTTCAGGCTCGACGCTCGCGCTTCACGGAGGAGCACTGCAAGGAC 126
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerIle 700
Db 127 ATCGGTTCTGCGCTCGCGGCCAACCCCTCTCCAGCGAGCTCTCCCTCCGACCAACGAG 186
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 187 CTGGCGATGCGGGTGCACCTGGTGTGAGGCGCTGACAGCCCTGACACCTGCAAGATC 246
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGlyValGlnHisLeuTrpArg 739
Db 247 CAGAAGCTCAGCTGACAGAACTGCTCCTGACCGAGCGGGCTGCGGGGTCTGCCCGAGC 306
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGlu 759
Db 307 ACGTGTGCTCTCCCTGCCACCTCGGGAGCTGATCTCAGCGCAACCCCACTGGGGAC 366
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 367 GCCGCGCTGCGGCTGCTGTGAGGGCTCTGACCCCGGAGTCCACCTGAGAAAGCTG 426
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 427 CAGTTGGAGTACTGCGGCTGACGCGCGCCAGCTGCGAGCCCTCGGCTCGGTCTCAGG 486
Qy 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 487 GCCACGCGGGCTTGAAGGAGCTCACGGTGAAGCAACACACATCGCGAGCGCGGCC 546
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 547 CGGGTGTGGGCGAGGCTCTGGCAGACTCTCCCTGCGAGCTGGAGCGCTCAGGCTGAG 606
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859

```

```

Db 607 AACTGCGGTCTCACGCCAGCAAACTGTGCGGAATTGTGGCTCTCCAGGCC 666
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 667 TCGCTGAGGAGCTTGACCTGGGCGCAACCGGGCTGGGCGACGCGGCATAGCCGAGCTG 726
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 727 TGCCCCGGGCTTGTAGCCCGGCTCCCGCTCAAGACCTGTGGCTCTGGAGTGTGAC 786
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 787 ATCACCAGCTGCTGACAGACCTCTGCGGTGCTCTCCAGGCCAAGAGACCTCAAG 846
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuCysGluVal 939
Db 847 GAGCTCAGTCTGGCGGCAACAAGCTGGGCGAGAGGGCGCCGCTGTGTGCGAGGC 906
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeuThrAla 959
Db 907 CTGCTGAGCCCGGCTGCGAGCTGGAGTCCCTGTGGGTGAAGTCTTCAGCCTCAGCGCG 966
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 967 GCCTGCTGCGAGCAGCTCAGCTTGTGCTGACCCAGAACCAAGCATCTCTCGAACTTCAG 1026
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 1027 TTGAGCAGCAACAAGCTGGGTGACTCTGGCATCCAGGAGCTGTGCCAGGCCCTGAGCCAG 1086
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCys 1019
Db 1087 CCGGGCACCACATGCGGGTGCTCTGCTTGGGGACTGTGAGGTGACCAACAGCGGTGC 1146
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 1147 AGCAGCTCGCTCGCTGCTGCTGGCCCAACCGCAGCCTGCGAGAGCTGGACCTGAGCAAC 1206
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 1207 AACTGTGGGCGACCCGCGCGCTCTGCTGAGCTGTGGGAGCTGTGAGAGCGCGGTGC 1266
Qy 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 1267 GCCTGGAGCAGCTGGTCTGTACGACACCTACTGAGCGGAGGAGGTGGAGGACCGCTG 1326
Qy 1080 GluGluValGlnLeuLysPro-----ArgValVal 1090
Db 1327 CAGGCCCTGGAGGAGCAAGCCCGGCTGAGGGTCAATC 1365

```

RESULT 5

US-08-795-395-1
 ; Sequence 1, Application US/08795395
 ; Patent No. 5965399
 ; GENERAL INFORMATION:
 ; APPLICANT: CHATTERJEE, DEB K.
 ; APPLICANT: SHANDILYA, HARINI
 ; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
 ; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/795,395
 ; FILING DATE: 04-FEB-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/024,057
 ; FILING DATE: 16-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ESMOND, ROBERT W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0942.3440002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1371 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1368
 ; US-08-795-395-1

Alignment Scores:

Pred. No.: 2,44e-55 Length: 1371
 Score: 621.00 Matches: 151
 Percent Similarity: 50.11% Conservative: 86
 Best Local Similarity: 31.92% Mismatches: 210
 Query Match: 8.78% Indels: 26
 DB: 2 Gaps: 5

US-10-066-521-6 (1-1344) x US-08-795-395-1 (1-1371)

Qy 625 AsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIle 644
 Db 4 AACCTCGATATT-----CATTCGAGCAGCTGAGCGAGCC 39
 Qy 645 ArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValVal 664
 Db 40 CGG-----TGGACAGAGCTCTCGCGCTGCTC 66
 Qy 665 ProLeuTrp-----MetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
 Db 67 CAGCAGTATGAGGTGGTTCAGGCTCGAGCCTCGCGGCTCGAGAGGACACTGCAAGGAC 126
 Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerIle 700
 Db 127 ATCGGTCTGCGCTCCGGGCGCAACCTCTCCAGCAGCTCTGCTCCGACCAACAGCAG 186
 Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
 Db 187 CTGGCGATGCGCGGTGCACCTGGTGTGAGGCGCTCGAGGCGCCACCTGCAAGATC 246
 Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGlyValGlnHisLeuTrpArg 739
 Db 247 CAGAAGCTTCAGCTCGAGAACTGCTCCCTGACCGAGCGGGGCTCGCGGCTCGCCAGC 306
 Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLysGlu 759
 Db 307 ACGTGGCTCCCTGCCACCTCGGGAGCTGATCTCAGCGCAACCCACTGGGGGAC 366
 Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
 Db 367 GCCGCGCTGGCGTCTGTGAGGGGCTCTCGAGCGCCCTGAGCCCTGGAGAGCTG 426
 Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
 Db 427 CAGTTGAGTACTCCGCGCTGACCGCGCCGAGCTGCGAGCGCCCTGCGTCTCGAGG 486
 Qy 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819

Db 487 GCCACGCGCGCTTGAGAGGCTTCAGGTGAGCAACACGACATCGCGAGGCGCGGCC 546
 Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
 Db 547 CGGGTGTGGGCCAGGGTCTGGCAGACTCTGCTGCCAGCTGGAGAGCGCTCAGGCTGGAG 606
 Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
 Db 607 AACTGCGGTCTCAGCCAGCAACTGTCGGAATTTGTGGCTCTCCAGGCC 666
 Qy 860 SerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGlnGlyValAsnLeuLeu 879
 Db 667 TCCTGAGGAGCTTGACCTGGCAGCAACGGGCTGGGCGACCGCGGCATACCGGAGCTG 726
 Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
 Db 727 TGCCCCGGGTCTTGAGCCCCCGCTCCCGCTCAAGACCTGTGGCTCTGGGAGTGTGAC 786
 Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
 Db 787 ATCACCAGCCAGTGGCTGCAGAGACCTCTGCGGTCTCTCCAGGCCAAGGAGACCTGAAG 846
 Qy 920 HisLeuSerSerMetAsnProValGluAspAsnGlyValLysLeuCysGluVal 939
 Db 847 GAGCTCAGTCTGGCGGCAACAGCTGGGCGAGCGGCGCGCTGTGTGCGAGAGC 906
 Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
 Db 907 CTGCTCAGCGCGGCTGCGAGCTGGAGTCCCTGTGGGTGAAGTCTTGACGCTCAGCGCG 966
 Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
 Db 967 GCCTGCTGCCAGCAGCTCAGCTTGATGCTGACCCAGCAACAGCATCTCTGGAACCTCAG 1026
 Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
 Db 1027 TTGAGCAGCAACAGCTGGTGTACTTGGCATCCAGGAGCTGTGCCGCGCTGAGCCAG 1086
 Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
 Db 1087 CCGGGCACCACACTGCGGGTGTCTCTTTGGGAGCTGTGAGGTGACCAACAGCGGCTGC 1146
 Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
 Db 1147 AGCAGCTCTCGCTCGCTCTCTGCGCAACCGCAGCTCGAGAGCTGGACCTGAGCAAC 1206
 Qy 1040 AsnAsnPheSerProLysGlyMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
 Db 1207 AACTGTGGGCGACCGCGCGCTCTGAGCTGTGGGAGCTGTGGAGCGCTGGAGCGGCTGC 1266
 Qy 1060 AsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
 Db 1267 GCCTTGAGCAGCTGGTCTGTACAGACACTACTGGACGAGGAGGTGGAGGACCGCCTG 1326
 Qy 1080 GluLysValGlnLeuLysPro-----ArgValVal 1090
 Db 1327 CAGGCCCTGGAGGGAGCAAGCCCGGCTGAGGGTCAATC 1365

RESULT 6

; US-08-910-731-7
 ; Sequence 7, Application US/08910731
 ; Patent No. 5932440
 ; GENERAL INFORMATION:
 ; APPLICANT: SHANDILYA, HARINI
 ; APPLICANT: CHATTERJEE, DEB K.
 ; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 ; CITY: WASHINGTON
 ; STATE: D.C.

COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/910,731
 FILING DATE: (Herewith)
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/795,395
 FILING DATE: 04-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/794,546
 FILING DATE: 03-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/024,057
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: ESMOND, ROBERT W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0942.3440003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1371 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cDNA

US-08-910-731-7

Alignment Scores:

Pred. No.: 2,19e-54 Length: 1371
 Score: 612.00 Matches: 152
 Percent Similarity: 49.26% Conservative: 81
 Best Local Similarity: 32.14% Mismatches: 214
 Query Match: 8.65% Indels: 26
 DB: 2 Gaps: 5

US-10-066-521-6 (1-1344) x US-08-910-731-7 (1-1371)

```

Qy 625 AsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIle 644
Db 4 AACCTGGACATT-----CATTCGAGCAGCTGAGCGACGCC 39
Qy 645 ArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValVal 664
Db 40 CGG-----TGGACAGAGCTCCTCCGCTGCTC 66
Qy 665 ProLeuTrp-----MetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 67 CAGCAGTATGAGTGCTCAGGCTCGACCACTCGCGCCCTCAGGAGGAGCACTGCAAGGAC 126
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 127 ATCGGTTCTCGCCCTCCGGGCAACCCCTCCCTGACCGAGCTCTGCCTCCGACCAACGAG 186
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 187 CTGGCGGATGCGGGCTGCACCTGCTGTCGAGGCGCTGAGAGCCCACTGCAAGATC 246
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGlyValGlnHisLeuTrpArg 739
Db 247 CAGAAGCTCAGCTGCAGAACTGCTCCCTGACCGAGGGGGCTCGGGGCTCCTGCCAGC 306
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGlu 759
Db 307 ACGTGGCTCCCTGCCCAACGCTGGGGAGGTGTCATCTCAGCGACAACCCACTGGGGGAC 366

```

```

Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 367 GCCGGCTGGCTGCTCTGTGAGGGGCTCTGGAGCCCGCCAGTCCACCTGGAGAAGCTG 426
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 427 CAGTTGAGTACTGCGGCTGAGCGCGCCAGCTGCGAGCCCTCGGCTCGGCTCTCAGG 486
Qy 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 487 GCCACGGGGCTTGAAGGAGCTCACGCTGAGCAACACACATCGCGGAGCGCGGCC 546
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 547 CGGGTGTGGCCAGGGTCTGGCCGACTCTGCCCTGCCAGCTGGAGAGCTCAGGCTGGAG 606
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 607 AACTGCGGTCTCAGCCAGCCAACTGCAAGACCTGTGCGGAATTGTGGCTTCCAGGCC 666
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 667 TCGTGGGAGCTGGCCCTGGGCAGCAACAAGCTGGGTGATGTGGGCATGGCGGAGCTG 726
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 727 TGCCAGGGCTGTCCACCCAGCTCCAGGCTCAGGACCTGTGGATCTGGAGTGTGGC 786
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 787 ATCACTGCCAAGGGCTGCGGGATCTGTGCTCTCAGGGCCAAAGAGAGAGCTCAAG 846
Qy 920 HisLeuSerLeuSerMetAsnProValGluAsnGlyValLysLeuLeuCysGluVal 939
Db 847 GAGCTCAGCTGGCCGCAACGAGCTGGGGATGAGGGTGGCCGCTGTGTGTGAGACC 906
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeuThrAla 959
Db 907 CTGCTGGAACCTGGCTGCCAGCTGGAGTCTGCTGGGTGAAGTCTTCAGCTTCACAGCC 966
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 967 GCCTGTGCCCCCACTTTCAGCTCAGTCTGCTGGCCCAAGAGGTTTCTCTCTGGAGCTACAG 1026
Qy 980 LeuThrAspAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 1027 ATAAGCAACACAGCTGGAGGATGCGGGCTGGGGAGCTGTGCCAGGGCTTGGGCCAG 1086
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCys 1019
Db 1087 CCTGGCTCTGTGCTGGGTGCTCTGTTGGCCGACTGCGATGTGATGTGACAGCAGCTGC 1146
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 1147 AGCAGCTCCCGCAACCTGTTGGCCAAACACAGCTGCTGCTGAGCTGGACCTCAGCAAC 1206
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 1207 AACTGCTGGGGAGCGCGGCATCCTGCACTGTGGAGAGCGTCCGGCAGCCGGGCTGC 1266
Qy 1060 AsnLeuGlnIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu 1079
Db 1267 CTCTGGAGAGCTGGTCTGTACGACATTTACTGGTCTGAGGAGATGGAGGACCGGCTG 1326
Qy 1080 GluGluValGlnLeuLysPro-----ArgValVal 1090
Db 1327 CAGGCCCTGGAGAGGACAAGCCATCCTCCTGAGGGTCAATC 1365

```

RESULT 7

US-08-910-731-5
 ; Sequence 5, Application US/08910731
 ; Patent No. 5932440

; Sequence 1904, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1904
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1515)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1904

Alignment Scores:

Pred. No.:	5,448-38	Length:	1515
Score:	458.00	Matches:	149
Percent Similarity:	45.83%	Conservative:	87
Best Local Similarity:	28.93%	Mismatches:	221
Query Match:	6.47%	Indels:	60
DB:	4	Gaps:	13

US-10-066-521-6 (1-1344) x US-09-833-381-1904 (1-1515)

```
Qy 4 AspLysSerLeuThrPheSerSerTyrGly-----LeuGlnTrpCysLeuTyrGluLeu 21
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 60 GACAAAGTGGTGTCTTCGGCGCAGATGGGCTTCAACCTGCGAGGCTCTCTCGAGCAGCTC 119
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 22 AspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysLysSerSerGluSerThr 41
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 AGCCAGGATGAGTTCAGCAAGTTCAGATATCTGATCAGCAGCTTCTCTCCCTGGCACACGAG 179
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 42 ThrCysSerIleProGlnPheGluIleGluAlaSerValGluCysLeuAlaLeuLeu 61
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 180 CTCGAGAGATCCCCCAACAGAGGATAGCAAGGCTGATGGGAGCAACTGGTAGAAATC 239
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 62 LeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsn 81
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 240 CTCACCCACCAATTGACAGCTACTGGGTGGAGATGGCGAGCTCCAGGCTTTTGAAAG 299
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 82 MetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysLysIleSerGlnAla 101
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 300 ATGCACCGAATGATCTCTCTGAGAGAGCAAGGATGAAGTCAGAGAGCAGCTTTGAAA 359
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 102 MetGluGlnGlu-----GlyAlaThrAlaAlaGluThrGluGluGln 115
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 360 TCCTTTTAATAAAGGAAGCCCTCTATCATTAGGATATAACACGGAAGAACACCACTCTA 419
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 420 GACATCGACCAATTAAGAA-----CTCTCAAAACAAAAAAACAA--- 461
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMet---ThrLysPheAlaGluGlu 154
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 462 -----GACAAAGACAATAGGTGACAGGTATATATTGAAGACGAAGTTTC----- 503
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 155 GluAspValArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 504 -----CGGAGATCTGGAAGAGCTGGCTGAGATAGCAAGAGGATCCAGGTATG 554
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArg----- 188
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 555 GCTGAGATACAGATCTGATCCATTTAGCAACCCAGGGGTCTTCGGGGCCCTTC 614
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 189 -----ThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 206
```

```
Db 615 TCATACACGGTGGTGTGTATGTCTCTGAGGCGCTTGGGAAACACCGCTGGCCACAGAA 674
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 207 IleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 675 CTAATGTAGATGGGACAGGACAACCTCATCCAC---AAATTCAAATATGCGTTCTTAC 731
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 227 LeuProValArgGluMetGlnArgLysGluSerSerValThrGluPheIleSerArg 246
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 732 CTCAGCTGACGGAGCTCAGCGCTGGGCCCGTGCAGTGTTCGACAGCTGGTCTTCAGG 791
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 792 GACTGGCTGAATTCAGGATGACATTCACACATCTAGCCCAAGCACGGAATCTTG 851
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 267 PheIleIleAspGlyPheAspAspLeuGlySer-----ValLeuAsnAspThr 283
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 852 TTGCTGATTACGGCTTTGATGAGCTGGGAGCCGACCTGGGGCGCTGATCGAGGAC--- 908
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 284 LysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeu 303
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 909 ---ATCTCGGGGACTGGGAGAGAAGACCGCGTGCCTCTCTCTGGGGAGTTTCTG 965
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 304 ArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGlu 323
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 966 AACAGGCTGATTTACCAAGGCGCCCTGCTGTGTACACACGCGCCAGGCCCTCAGG 1025
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 324 LysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGlu 343
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1026 GACCTCGGATCTCGCGGAGGAGCCGATCTACATAGGCTGGAGGCTTCTCTGAGGAG 1085
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 344 GlnArgIleHisLeuLeuLeuArgGlyIleGlyGluHisGlnLysThrGlnGlyLeu 363
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1086 GACAAGAGGCGCTATTCTCTGAGACATTTGGAGACGAGGACCAAGCCATGCGTGCCTTT 1145
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 364 ArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySer 383
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1146 GAGCTAATGAGGACACGCGGCCCTGCTCCAGCTGGGCTCGGCCCGCGGCTGTGCTGG 1205
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 384 LeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsn 403
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1206 ATCGTGTGACAGCTCTGAAGCTGCAGATGAGAAAGGAGGAGACCGGCTCCCACTGC 1265
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 404 GlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro----- 420
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1266 CTCACCGCAGCGGGCTGTT-CTTGGCTTCTCTGAGCGGCTTCCCGCAGGCGCAGAC 1324
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 421 ---ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1325 NTGCGGGCGCGCTGCGACGCTGAGCCTCTGCGCGCGCAGGGCTGTG----- 1372
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1373 -----GGGCAAGTGTCCGTGTCCTGTCACCGAGGAGGACCTG 1405
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1406 GNAAGGCTCGGGTGCAGGAGTCCGACNTCCGNTGTTCTCTGAGCAGGAGANATCTCCG- 1464
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 480 ProAsp-----SerHisCysGluGluTyrTyrThrPhePhe 491
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1465 CCAGACAGAGTCTCCAAAGGCTGTCGCTTATCCANTCAGNTTC 1509
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

RESULT 9

US-09-833-381-1905
; Sequence 1905, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11

;; PRIOR APPLICATION NUMBER: 09/516,448
;; PRIOR FILING DATE: 2000-02-29
;; NUMBER OF SEQ ID NOS: 2050
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1905
;; LENGTH: 1515
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(1515)
;; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1905

Alignment Scores:
Pred. No.: 5,44e-38 Length: 1515
Score: 458.00 Matches: 149
Percent Similarity: 45.83% Conservative: 87
Best Local Similarity: 28.93% Mismatches: 221
Query Match: 6.47% Indels: 60
DB: 4 Gaps: 13

US-10-066-521-6 (1-1344) x US-09-833-381-1905 (1-1515)

Qy 4 AsplysSerLeuThrPheSerSerTyrGly-----LeuGlnTrpCysLeuTyrGluLeu 21
Db 60 GACAGATGGTCTTCGGCCAGATGGCTTCAACCTGCAGGCTCTCTGGACGACCTC 119
Qy 22 AsplysGluGluPheGlnThrPheLysGluLeuLeuTyrLysSerSerGluSerThr 41
Db 120 AGCCAGGATGATTCAGCAAGTTCAGATCTGTATCAGCAGCTTCTCCCTGGCACAG 179
Qy 42 ThrCysSerIleProGlnPheGluLeuGlnAlaSerValGluCysLeuAlaLeuLeu 61
Db 180 CTCAGAGATCCCCCAAGAGAGGTAGCAAGGCTGATGGGAAGCAACTGGTAGAATC 239
Qy 62 LeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsn 81
Db 240 CTCACCCCAATTGTGACAGTACTGGTGGAGATGGGAGCTTCCAGGCTCTTGAAGAAG 299
Qy 82 MetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysIleSerGlnAla 101
Db 300 ATGCACCAATGGATCTGTCTGAGAGAGCAAGGATGAAGTCAGAGAAGCAGCTTTGAAA 359
Qy 102 MetGluGlnGlu-----GlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 TCCTTTAATAAAGGAAGCCCTCTATCATATTAGGATACACGGAAGAACACCACTCTA 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135
Db 420 GACATCGACCAATTAAGAA-----CTCTTCAAAACAAACAA----- 461
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMet---ThrLysPheAlaGluGlu 154
Db 462 -----GACAAAGCAATAGGTGCGAGGTATATATTGAAGACGAAGTTC----- 503
Qy 155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 504 -----CGGAGATGTGTGAAGAGTGGCGCTGGAGATAGCAAGAGGTCCAGGTTATG 554
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArg----- 188
Db 555 GCTGAGATACAGATGCTGATCCCATTTAGCAACCCAGGGTGTCTCCGGGCCCTTC 614
Qy 189 -----ThrValValLeuHisGlyLysSerGlyLeuGlyLysSerAlaLeuAlaArgArg 206
Db 615 TCATACAGGTGGTGTGTATGTGTCTCGAGGCTTGGGAAACCAACGCTGGCCAGAAA 674
Qy 207 IleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 675 CTAATGTAGATGGGAGAGCAACCTCATCCAC---AAATTCAATATGCGTTCTAC 731
Qy 227 LeuProValArgGluMetGlnArgLysGluSerSerValThrGluPheIleSerArg 246

Db 732 CTCAGTCGAGGAGCTCAGCCCTGGGCCCTGGAGTTTTCAGAGCTGGTTCAGG 791
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 792 GACTGGCTGAATTGACAGATGATTCACACATCTTAGCCCCAAGCAGCAGAAATCTTG 851
Qy 267 PheIleIleAspGlyPheAspLeuGlySer-----ValLeuAsnAsnAspThr 283
Db 852 TTCGTGATTGACGGCTTTGATGAGCTGGAGCCGACCTGGGGCGCTGATCGAGGAC-- 908
Qy 284 LysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeu 303
Db 909 ---ATCTCGGGGACTGGGAGAAAGAACCCCGTCCCGTCTCTCTGGGAGTTTGTCTG 965
Qy 304 ArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGlu 323
Db 966 AACAGGCTGATGTTACCAAGCGCCCTGCTGTCCACACGCGGCCAGGGCCCTGAGG 1025
Qy 324 LysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyLysSerGlyGlu 343
Db 1026 GACCTCCGATCTCGCGAGGAGCGCATCATTAAGGTTGGAGGGCTTCTCTGGAGGAG 1085
Qy 344 GlnArgIleHisLeuLeuGluArgGlyLysGlyLysGlnLysThrGlnGlyLeu 363
Db 1086 GACAAGAGGCTTATTTCTGAGACACTTTGGAGACGAGACCAAGCCATGCGTGCCTTT 1145
Qy 364 ArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySer 383
Db 1146 GAGCTAATGAGGAGCAACGCGNCCCTGTTCCAGCTGGGCTCGGCCCGCGGTGTCTGG 1205
Qy 384 LeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsn 403
Db 1206 ATCGTGTGACGACTCTGAAGCTGCAGATGNAAAGGAGGAGCCCGTCCCGACCTGC 1265
Qy 404 GlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro----- 420
Db 1266 CTCACCCGCGCGGGCTGTT--CTTGGCTTCTCTCGACCGCGGTTCGCCGAGCGCACAG 1324
Qy 421 ---ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
Db 1325 NTGCGGGCGCGCTGCGAGCTGAGGCTCTTGGCGCGCAGGGCTGTG----- 1372
Qy 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
Db 1373 -----GGGCAGATGTCCGTGTNCCACCGAGAGGACCTG 1405
Qy 460 MetValGlnGlyLeuGlySerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1406 GNAAGCTCGGGTGCAGGAGTCCGACNTCCGNTGTTCTCTGACGAGGAGATCTCCG- 1464
Qy 480 ProAsp-----SerHisCysGluGluTyrTyrThrPhePhe 491
Db 1465 CCAGACAGAGTCTCCAAAGGCTGTNGCCTTATCCANTCAGNTTC 1509

RESULT 10

US-09-833-381-1976
; Sequence 1976, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1976
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens


```
Db 2388 CTGGCAACAACAACCTCAATGACTATCGCTGCAG-----GAGCTGACCGCTTGTCTT 2441
Qy 668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db 2442 AGCGCT----- 2447
Qy 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLys 707
Db 2448 -----CTCAGCGTTATCAGACTCAGCGTCAACACAGATCACCAGACCGCGGGTGAAG 2498
Qy 708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet---PheArg 726
Db 2499 GTGCTATGTAGGAAGT-----ACCAAGTATAGATGCTGAGCGTCTCTGGGTTTATAC 2552
Qy 727 AsnAlaGlnIleThrPro---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
Db 2553 AACCAACAGATACTGATATCGGACCCAGGTATGTGGCCCAATCTCTGGATGATGAGCA 2612
Qy 746 AsnLeuArgSerLeuAsnLeuGlyGlyThrHisLysLysGluGluAspValArgMetAla 765
Db 2613 GGCCTCAAGCACCTTAAACTAGGGAACCAAGATAACAAGTGAAGGCGGGAAGTGTGTG 2672
Qy 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGly 785
Db 2673 GCTTGGCTGTGAGAAC----- 2690
Qy 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
Db 2691 -----AGCACCTCCATCGTT 2705
Qy 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825
Db 2706 GATGTTGGGATGTGGGTAAATCAGATTGCAGACGAAGGGCAAGGCCCTTCGCAGAGCA 2765
Qy 826 LeuArgValSerGlnCysAlaLeuGlnLysIleLeuLeuGluAspCysGlyIleThrAla 845
Db 2766 TTGAAG---GACCACCCCGCTGACCACTCTCAGTCTTGATCTCAATGGCATCTCCG 2822
Qy 846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
Db 2823 GAGGAGGAGAGAGCTTGCAGCGCTTGAAGCAGAACACACACACACATGACGATATCTGG 2882
Qy 866 LeuSerAsnSerLeuGlyAsnGlyValAsnLeuLeuCysArgSerMetArgLeu 885
Db 2883 CTGACCAAAATGAATTAATGATGAGTCTCAGAGTCTTGGCTGAGATGCTCAGAGTG 2942
Qy 886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
Db 2943 ---AACCCAGCGCTACGCGATTTATGGCTGATCCAGAAATCCGATCAGCCAGCGGACA 2999
Qy 906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMet 925
Db 3000 CGCAGCTGGCGGGCACTGCAGNAGAACACAGCCATAACAGAGATTTGTCTCAATGGA 3059
Qy 926 AsnProValGlu-----AspAsnGlyValLysLeuLeuCysGlu 938
Db 3060 AACTGATTAAAGCCGAGGAGGCCAAAGTCTTCGAGAAATGAGAGAGAAATCATCTGCTT- 3118
Qy 939 ValMetArgGluProSerCysHisLeuGlnAspLeu-----GluLeuValLysCysHis 956
Db 3119 CTGACGGACGCTCTGGG-----CAGGATCTTTGTCTAGGTTGCTCTCAGTCAC 3169
Qy 957 LeuThrAlaAlaCys-----CysGluSerLeuSerCysVal 968
Db 3170 AG-ACAGCACTGTGAGTCAGCAGGAGTAGCAGGATGCTGTGCAGCGCTGCAGCAAGGTG 3228
Qy 969 IleSerArgSerArgHisLeuLysSer 977
Db 3229 CCTGTGAGGAGCCACACCTCCACAGT 3255
```

RESULT 12

US-09-207-359B-42

; Sequence 42, Application US/09207359B

```
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; CURRENT FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 4141
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (261)...(3119)
US-09-207-359B-42
```

```
Alignment Scores:
Pred. No.: 7,36e-29 Length: 4141
Score: 379,50 Matches: 253
Percent Similarity: 37,01% Conservative: 150
Best Local Similarity: 23,23% Mismatches: 416
Query Match: 5,36% Indels: 273
DB: 4 Gaps: 46
```

US-10-066-521-6 (1-1344) x US-09-207-359B-42 (1-4141)

```
Qy 6 SerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGlu 25
Db 451 AGCTGACAAAGGTCGAAAGATCCTTGACCTGTGC----- 486
Qy 26 PheGlnThrPheLysGluLeuLeuLysSerSerGluSerThrThrCysSerIle 45
Db 487 -----AGACCAAGGCGAGGAGGTCTGAGTTCTTCCTCTAGCTGTGAG- 533
Qy 46 ProGlnPheGluLeuGlnAlaAsnValGluCysLeuAlaLeuLeuHisGluTyr 65
Db 534 -----CAGCTGGAGGATGCTTACGTGAGC---CTCAGGCTG----- 566
Qy 66 TyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsnMetAsnLeuArg 85
Db 567 -----TGGCTCTCAGAAATTTGGCTTCTCCCTTCCAGCTCATTTCCG 608
Qy 86 ThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGlnAlaMetGluGlnGlu 105
Db 609 ACCAAACTATCGTCAATACCTGACCAGTAAGCAGGTATACCCACAGCTGCGACACAA 668
Qy 106 GlyAlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAlaMetGluGlnGly 125
Db 668 ----- 668
Qy 126 AlaThrAlaAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer 145
Db 669 -----CTGGCGCGGAC-----TCC 683
Qy 146 HisValMetThrLysPheAlaGluGluAspValArgArgSerPheGluAsnThrAla 165
Db 684 AAGTTCACTGTCTGTCTAGCGCCGAGAGGAGGAGCTG-----CTGCTGGAGGAGACC- 734
Qy 166 AlaAspTrpProGluMetGlnThrLeuAlaGly-----AlaPheAspSerAspArgTrp 183
Db 735 -----TATATGGACACACTCATGGGCTGTGTAGGCTTCAACATGAAGAACTG 782
Qy 184 GlyPhe-----Arg 186
Db 783 GGCAGCGCTAGGAGCGCTGGATTGCTCTGACCAACAGTACGGGCGTCTCAACAGAGCAT 842
```

QY 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 206
DB 843 GCGAGACTGCTTCGTTTCGGGACGCGGAGTGGGCAAGTCACTGCTGCTCAGAGG 902
QY 207 IleValLeuCysTrpAlaGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
DB 903 TTGCAGAGCCTCTGGGCGTCAGGCAGGTTG--ACCTCCACAGCCAAATCTCTTCCAC 959
QY 227 LeuProValArgGluMetGlnArgLysLysGluSer-----SerValThrGluPhe 243
DB 960 TTCGCTGCGCCATGTTTCAGCTGCTTCAAGAGAGCAGCATGCTGAGCTCTGAGAGCTG 1019
QY 244 IleSerArgGluTrp-----ProAspSerGlnAlaProValThrGluLeuMet 259
DB 1020 CTCCTCAGCATTTCTGCTACCGGAGCAGACCCCGAGGAGGTGTTCTCTCTTCTGCTG 1079
QY 260 SerArgProGluArgLeuPheIleAspGlyPheAspLeuGlySer----- 277
DB 1080 CGCTTTCCCCACACAGCGCTCTTCACTTTTCAGCGCTGGATGAGCTGCACTCAGATTC 1139
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThr 297
DB 1140 GACCTGAGCGCGTCCCGATAGCTGCTGCCCTGG---GAGCCGGCTCACCTCTCGCT 1196
QY 298 LeuIleArgSerLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB 1197 CTGCTGGCTAACTTAACTGAGGAGCTGCTCAAGGTCGCGCAAAATGCTCACCTGCT 1256
QY 318 Arg-----AspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
DB 1257 GGCACAGCGTGGAGTCCCGCCAGCTCTCGCCAAAAGTG----- 1301
QY 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeu----- 348
DB 1302 ---CTGCTCCGGGCTTCTCCCAAGT-----CACCTCGCGCTATGCCCGCGG 1349
QY 349 ---LeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMet 367
DB 1350 ATGTTCCCGAGCGCAGCCGACGAGCATCTG-----CTGCAGCAGCTGGAT 1397
QY 368 AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal 387
DB 1398 GCCAACCCCACTCTGCAGCTGTCGGGGTGGCGCTCTTCTTGTGATCATCTCCGT 1457
QY 388 AlaLeuGln----LeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
DB 1458 TGTTTCCAGCATTTCCAGACGCTTTCGAGGCTCTCTTCACAGTTGCGGAGCTGTGCT 1517
QY 407 ThrGlyLeuHisAlaAlaPheVal-----PheHisGlnLeuThrPro 420
DB 1518 GTGACCTGACCGATGCTTTCTGCTGCTCACTGAGTGCATCTGAACAGCCGCGACGCC 1577
QY 421 ArgGlyValValArgCysLeuAsnLeuGluGluArgValVal----- 435
DB 1578 AGCAGCTGGTGCAGCGCAACACGCGACCGCGCGGAAACCTACCTGCGAGCTGGCGC 1637
QY 436 ---LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
DB 1638 ACGTGCATGCGCTGGGAGAGCTGCTCACCGAGCCCGCAGAGACCTCTTTGTGTTT 1697
QY 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
DB 1698 GCCCAGGAGGAGTGCAGGCTCAAGCTGCAGAGAGAGATCTGCAGCTGGGCTTCTG 1757
QY 475 MetAsnIleLeuLeuProAspSerHisCysGlu-----GluTyrTyrThrPhePheHis 492
DB 1758 CGG-----GCTTTGCCGATGTGGGCCCTGAGCAGGCGCCAGCTTACGAATTTTCCAC 1811
QY 493 LeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGlu 512
DB 1812 CTTACGCTCCAGGCTTCTTCCCGCCCTTCTCTGTCGATGAGATGCAAGAGTGAGCACC 1871
QY 513 ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla 532

DB 1872 CGGAGTTGCTAGGTTCTTTCAGAGATGACGCTCTCTGGA---GAGGCAACAAGCTCG 1928
QY 533 GlyPheHisIle----- 536
DB 1929 TCCTGCCATTTCTTCTTCTCTTCCAGTGCTGGGGGCAAGACCGGTTGGGCGCT 1988
QY 537 -----HisSerLeuTrpMetLysArgPheLeuPheGlyLeuVal 549
DB 1989 GATCCTTTTCAGGAACAAAGATCACTTCCAGTTTCCCAACCTCTTCTGTCGGGCTACTG 2048
QY 550 SerGluAspValArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVal 569
DB 2049 GCCAAAGCCGACAGAACTCTTTCGCGAGCTG-----GTGCCAAGGCTATC 2096
QY 570 -----LysGlnLysLeuLeuHisTrpVal-----SerLeu 579
DB 2097 CTGAGGAGGAGGCAAGGCCCTG---TGGGCTCACCTGTTTGTAGCTGCGCTCTCTAC 2153
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGly-----Asp 591
DB 2154 TTGAAGAGCCTACTCTCGGCTCCAGTCTGGAGGCTTTAAACAGGTGCTGCTGCCACCA 2213
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
DB 2214 TTCTGTGGATGCTGCGCTGCTATGACGACGAGCAGAGGTGGGCGCTCGCC 2273
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIle-----AsnGlnAsnLeuAsp 627
DB 2274 GCCAGGGCATCAGTGGGACTACCTCAAGCTGGCTTTTGCAACGCTTGTCTCGGAC 2333
QY 628 LeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAsp 647
DB 2334 TGCAGCGCTGCTCTTCTGCTGCTATC-----TTCACAGGAGCTGGCGCTAGAC 2387
QY 648 ValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrp 667
DB 2388 CTGCACACACAACTCAATGACTATGCGCTGAG-----GAGCTGCGAGCTTGTCTT 2441
QY 668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
DB 2442 AGCCGT----- 2447
QY 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLys 707
DB 2448 -----CTCAGGTTATCAGCTCAGCTCAACAGATCACCACAGCGGGGTGAAG 2498
QY 708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet---PheArg 726
DB 2499 GTGCTATGTGAGAACTG-----ACCAAGTATAGATCGTGACGTTCTGGGTTTATAC 2552
QY 727 AsnAlaGlnIleThrPro---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
DB 2553 AACCAACAGATACTCATATCGGAGCCAGGTATGTCGCCCAATCTCTGGATGAATGAGA 2612
QY 746 AsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGluAspValArgMetAla 765
DB 2613 GGCTCAAGCACCTTAACTAGGGAATAACAGAAATACAGTGAGGGCGGAGGTGTGTG 2672
QY 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGly 785
DB 2673 GCTTTGCTGTGAAGAAC----- 2690
QY 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
DB 2691 -----AGCACCTCTCATCTG 2705
QY 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825
DB 2706 GATTTGGGATGGGGTATCAGTTGGAGACGAAGGGGCAAGGCCCTTTCGACAGAGCA 2765
QY 826 LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAla 845

Db 2766 TTGAAG--GACCACCCAGCGCTGACCACTCTCAGTCTTGCAATCAATGGCATCTCTCG 2822
Qy 846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
Db 2823 GAGGAGGAGAGAGCTTGGCGGCGCTGAGCAGACACACACTGACAGATTAATCTGG 2882
Qy 866 LeuSerAsnAsnSerLeuGlyAsnGlyValAsnLeuLeuCysArgSerMetArgLeu 885
Db 2883 CTGACCAAAATGAACCTTAATGATGAGTCTGAGAGTCTGCTGAGATCTGAGAGTG 2942
Qy 886 ProHisCysSerLeuGlnArgLeuMetLeuAnGlnCysHisLeuAspThrAlaGlyCys 905
Db 2943 ---AACCCAGCTACCGCATTTATGGCTGATCCAGATCCATCCACACCAAGGGACA 2999
Qy 906 GlyPheLeuAlaLeuMetGlyAsnSerTTPLeuThrHisLeuSerLeuSerMet 925
Db 3000 GCGAGCTGGCGAGGCGACTGCAGAGAACAACAGCCATTAACAGATTTGTCTCAATGGA 3059
Qy 926 AsnProValGlu-----AspAsnGlyValLeuLeuLeuCysGlu 938
Db 3060 AACTTGATTAGCCCGGAGGAGCCAAAGTCTTCGAGATGAGAGAGAAATCATCTGCTT- 3118
Qy 939 ValMetArgGluProSerCysHisLeuGlnAspLeu-----GluLeuValLysCysHis 956
Db 3119 CTGACGAGCGCTCTCTGG-----CAGGATCTTTGCTCTAGGTGCTCTCAGTCAC 3169
Qy 957 LeuThrAlaAlaCys-----CysGluSerLeuSerCysVal 968
Db 3170 AG-ACAGCACTGTCAGTCAGCAGCGGTAGCAGGATGCTGTCAGCGCTGCAGCAAGGTG 3228
Qy 969 IleSerArgSerArgHisLeuLysSer 977
Db 3229 CCGTGCAGGAGCCACACCTCCACAGT 3255

RESULT 13

US-09-340-620A-42
; Sequence 42, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US/09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US/09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US/09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US/09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 4141
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (261)....(3119)
US-09-340-620A-42

Alignment Scores:
Pred. No.: 7,36e-29 Length: 4141
Score: 379.50 Matches: 253
Percent Similarity: 37.01% Conservative: 150
Best Local Similarity: 23.23% Mismatches: 416
Query Match: 5.36% Indels: 273
DB: 4 Gaps: 46

US-10-066-521-6 (1-1344) x US-09-340-620A-42 (1-4141)

Qy 6 SerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGlu 25
Db 451 AGCCTGACAGGTCGCAAGATCCTTGACCTGTC----- 486
Qy 26 PheGlnThrPheLysGluLeuLysLysSerSerGluSerThrThrCysSerIle 45
Db 487 -----AGAGCAAGGCGAGGAGTGTCTGAGTTCCTTCCTCTAGTCTGTCAG- 533
Qy 46 ProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeuLeuHisGluTyr 65
Db 534 -----CAGCTGGAGGATGCTTACGTGGAC--CTCAGGCTG----- 566
Qy 66 TyrGlyAlaSerLeuAlaTTPAlaThrSerIleSerIlePheGluAsnMetAsnLeuArg 85
Db 567 -----TGGCTCTCAGAAATGGCTTCTCCCTTCTCCAGCTCATTCGG 608
Qy 86 ThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGlnAlaMetGluGlnGlu 105
Db 609 ACCAAACTATCTCAATCTGACCCAGTAAGCAGGTATACCCACACAGCTGCCACACAA 668
Qy 106 GlyAlaThrAlaAlaGluThrGluGlnGlnIleSerGlnAlaMetGluGlnGluGly 125
Db 668 ----- 668
Qy 126 AlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer 145
Db 669 -----CTGGGCGCGCAG-----TCC 683
Qy 146 HisValMetThrLysPheAlaGluGluAspValArgSerPheGluAsnThrAla 165
Db 684 AAGTTTCATGCTGTGCTACGCCAGAGAGGAGGAGCCTG-----CTGCTGGAGGAGACC- 734
Qy 166 AlaAspTrpProGluMetGlnThrLeuAlaGly-----AlaPheAspSerAspArgTyr 183
Db 735 -----TATATGGACACACTCATCGGGCTGTGTAGCTTCAACATGAAGAACCTG 782
Qy 184 GlyPhe-----Arg 186
Db 783 GGCAGCCTAGGAGCGCTGGATTGCTCTGGACCCACACAGTACGGGCGCTCTCAACAGCAT 842
Qy 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 206
Db 843 GCGGAGACTGTCTTCGTGTTCCGGGAGCGCGGAGTGGCAAGTCCATGCTGCTGCAGAGG 902
Qy 207 IleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 903 TTGAGAGCCTCTGGGCGCTCAGGAGGTTG---ACCTCCACAGCCAAATCTCTTCCAC 959
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSer-----SerValThrGluPhe 243
Db 960 TTCCGCTGCCGATGTTTCAGCTGCTTTCAGAGGAGCGACATGCTGAGTCTGACGAGCCTG 1019
Qy 244 IleSerArgGluTrp-----ProAspSerGlnAlaProValThrGluIleMet 259
Db 1020 CTCTTCAAGCATTTCTCTCTACCCCGACAGCAGCCCGAGGAGGTGTTCTCTCTTCTGCTG 1079
Qy 260 SerArgProGluArgLeuPheIleAspGlyPheAspLeuGlySer----- 277
Db 1080 CGCTTTCCACACAGCGCTCTTTCATCTTTGACGCGCTGGATGAGTGCATCAGACTTC 1139
Qy 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 1140 GACCTGAGCGCGTGGCGGATAGCTGTGCGCCCTGG---GAGCGGCTCACCTCTGCTGTC 1196
Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1197 CTGCTGGCTAACCTCTCTAAGTGGAGGCTGCTCAAGGGTCCCGCAATGTCTCAGTCT 1256
Qy 318 Arg-----AspValGlyThrGluLysLysSerGluValValSerProArgTyr 334
Db 1257 CGCAGCGCTGGAGGTCCTCCCGCAGCTCTCTCGCAAAAGGTG----- 1301
Qy 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeu----- 348

RESULT 14

US-09-865-364-42
 ; Sequence 42, Application US/09865364
 ; Patent No. 6613521
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-112001
 ; CURRENT APPLICATION NUMBER: US/09/865,364
 ; CURRENT FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: US 09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 42
 ; LENGTH: 4141
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (261)...(3119)
 US-09-865-364-42

Alignment Scores:

Pred. No.: 7,366-29 Length: 4141
 Score: 379.50 Matches: 253
 Percent Similarity: 37.01% Conservative: 150
 Best Local Similarity: 23.23% Mismatches: 416
 Query Match: 5.36% Indels: 273
 Ds: 4 Gaps: 46

US-10-066-521-6 (1-1344) x US-09-865-364-42 (1-4141)

Db	735	-----TATATGGACACACTCATGGGGCTGGTAGGCTTCAACAATGAACCTG	782
Qy	184	GlYPhe-----	186
Db	783	GGAGCCTAGAGCGCTGGATTGCTCTGGACACAGTAGCGGCGTCTCAAGACGAT	842
Qy	187	ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg	206
Db	843	GGCGAGACTGCTTCGTGGTTCCGGAGCGGGAGTGGCAAGTCCATGCTGCTCAGAGG	902
Qy	207	IleValLeuCysTrpAlaGlnGlyGlyLeuTrpGlnGlyMetPheSerTyrValPhe	226
Db	903	TTGCAGAGCTCTGGCGCTCAGGAGGTG---ACCTCCACAGCAAAATCTTCTTCCAC	959
Qy	227	LeuProValArgGluMetGlnArgLysLysGluSer-----SerValThrGluPhe	243
Db	960	TTCCGCTGCCGATGTTACGCTGCTTCAAGAGAGCGACATGCTGAGTCTGCAGGACTG	1019
Qy	244	IleSerArgGluTrp-----ProAspSerGlnAlaProValThrGluIleMet	259
Db	1020	CTCTTCAAGCATTTCTGTACTACCGAGCAGGACCCGAGGAGGTGTTCTCTTCTTCTG	1079
Qy	260	SerArgProGluArgLeuPheIleAspGlyPheAspPheGlySer-----	277
Db	1080	CGCTTTCCACACACAGCGCTCTTCACTTTTGCAGCGCTGATGAGTGCATCAGACTTC	1139
Qy	278	ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr	297
Db	1140	GACCTGAGCGCGTGGCGGATAGTGTGCCCTGG---GAGCGGCTCACCTCTCGTGC	1196
Qy	298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	1197	CTGCTGGCTAACTCTTAAGTGGGAGGCTGCTCAAGGGTCCGCGCAAAATGCTCAGTCT	1256
Qy	318	Arg-----AspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr	334
Db	1257	CGCACAGCGCTGAGGTGCCCGCCAGCTCTCGCAAAAGGTG-----	1301
Qy	335	LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeu-----	348
Db	1302	---CTGCTCGGGCTTCTCCCAAGT-----CACCTGGCGCTATGCCCGCGG	1349
Qy	349	---LeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuAlaIleMet	367
Db	1350	ATGTTCCCGCGCGCACAGCGAGGACATCTG-----CTGCAGCAGCTGGAT	1397
Qy	368	AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal	387
Db	1398	GCCAAACCCCAACCTGTCAGCCTGTGGGGTGCCGCTCTTCTGTTGGATCATCTCCGT	1457
Qy	388	AlaLeuGln---LeuGlnAspValValGlyLysSerValAlaProPheAsnGlnThrLeu	406
Db	1458	TGTTTCAGCAGCTTCCAGCGCTTTCGAGGGCTCTCTTACAGTTCGCGGCTGTGCT	1517
Qy	407	ThrGlyLeuHisAlaAlaPheVal-----PheHisGlnLeuThrPro	420
Db	1518	GTGACCTGACCGATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1577
Qy	421	ArgGlyValValArgCysLysLeuAsnLeuGluGluArgValVal-----	435
Db	1578	AGCAGCGCTGTGCGCGCAACACCGCGCGAGCCCGGCAAAACCTTACGTGCGGCTGCGC	1637
Qy	436	---LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe	454
Db	1638	ACGTGATCGCTGGAGAGGTGGCTTCCCGAGGACCCGAGGACCTCTTGTGTTT	1697
Qy	455	AspGlyAspLeuMetValGlnGlyLeuGlySerGluLeuArgAlaLeuPheHis	474
Db	1698	GGCCAGGAGGAGTGCAGGCGTCCAGCTCAGAGGAGATCTGCAGTGGGCTTCTCTG	1757
Qy	475	MetAsnIleLeuLeuProAspSerHisCysGlu-----GluTyrTyrThrPhePheHis	492
Db	1758	CGG-----GCTTTGCCGATGTGGGCCCTGAGCGGCGGCGGCTTACGAATTTTCCAC	1811

```
QY 493 LeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGlu 512
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1812 CTTAGCTCCAGGCTCTTCTCACCAGCTTCTCTCGTAGCAGATGACAAAGTAGGACACC 1871
QY 513 ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla 532
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1872 CGGAGTGTCTGAGGTTCTTTGAGATGGAGCTCTCTGGA---GAGGCAACAGCTCG 1928
QY 533 GlyPheHisIle----- 536
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1929 TCCTGCCATTCTCTCTCTCCATCCAGTCCCTGGGCGGCAGAAAGCCGTTGGGCCCT 1988
QY 537 -----HisSerLeuTrpMetLysArgPheLeuPheGluVal 549
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1989 GATCCTTTCCAGGAACAAGATCACTCCAGTTTACCAACCTCTCTCGTGGCGGCTACTG 2048
QY 550 SerGluAspValArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVal 569
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2049 GCCAAGCCCGACAGAACTCTTCCGAGCTG-----GTGCCCAAGGCTATC 2096
QY 570 -----LysGlnLysLeuLeuHisTrpVal-----SerLeu 579
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2097 CTGAGGAGGAAGCGAAGGCCCTG---TGGGCTCACCTGTTTGTAGCTCGCGCTCCTAC 2153
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGly-----Asp 591
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2154 TTGAAGAGCTTACTCTGGGTCCAGTCTGGAGGCTTTAACAGGTGCATGCCATGCCACCA 2213
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2214 TTCCTGTGGTGTGGCTGCTGATATGAGAGCGAGAGCCAGAGGTGGGGCGCTCGCC 2273
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIle-----AsnGlnAsnLeuAsp 627
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2274 GCCAGGGGCATCAGTGGCGGACTACTCAAGTGGCGCTTTTGCAACGCTTGCTCTGCGGAC 2333
QY 628 LeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAsp 647
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2334 TGCAGCGCCCTGTCTTCTGCTCGATCAC-----TTCCACAGGCGAGCTGGCGCTAGAC 2387
QY 648 ValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrp 667
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2388 CTGGACAACAACCACTCAATGACTATGGCGTGAG-----GAGCTGCAGCCTTGCTTT 2441
QY 668 MetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2442 AGCCGT----- 2447
QY 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLys 707
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2448 -----CTCAGGTATCAGACTCAGCTCAACAGATCACCGACACGGGGGTGAAG 2498
QY 708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet---PheArg 726
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2499 GTGCTATGTAGGAACGTG-----ACCAAGTATAAGATCTGACGTTCTCTGGGTTTATAC 2552
QY 727 AsnAlaGlnIleThrPro---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2553 AACCAACAGATAAATGATATCGAGGCCAGGTATGGCCCAATCTCTGGATGAATGCAGA 2612
QY 746 AsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAla 765
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2613 GGCCTCAAGCACTTAACTAGGGAACAAAGATAAACAAGTGGCGGCGGAAGTGTGTG 2672
QY 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGly 785
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2673 GCTTTGGCTGTGAAGAAC----- 2690
QY 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2691 -----AGCACTTCATCGTT 2705
```

```
QY 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2706 GATGTTGGGATGTGGGTATATCAGATTGGAGACGAAGGGGCAAAAGGCTTTCGACAGGCA 2765
QY 826 LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAla 845
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2766 TTGAAG---GACCACCCAGCCTGACCACCTCTCAGCTTTCATTCATGSCATCTCTCCG 2822
QY 846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2823 GAGGAGGGAAGAGCCTTTCGCGAGGCCCTGAAGCAGACACACCACTGCACAGTAATCTGG 2882
QY 866 LeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeu 885
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2883 CTGACCAAAATGAATTAATGATGAGTCTGCAGAGTGTCTCGCTGAGATGCTGAGAGTG 2942
QY 886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2943 ---AACCAGACGCTACGGCATTTATGGCTGATCCAGATCGCATCACAGCAAGGGGACA 2999
QY 906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMet 925
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3000 GCGCAGCTGGCGAGGCGACTGCAGAAAGAACACAGCCATAACAGAGATTGTCTCAATGGA 3059
QY 926 AsnProValGlu-----AspAsnGlyValLysLeuLeuCysGlu 938
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3060 AACTGATTAAAGCCCGAGGAGGCCAAAGTCTTCGAGATGAGAGAGAAATCATCTGCTT- 3118
QY 939 ValMetArgGluProSerCysHisLeuGlnAspLeu-----GluLeuValLysCysHis 956
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3119 CTGACGAGCGCTCTCTGG-----CAGGATCTTTGCTAGTGTCTCTCAGTCAC 3169
QY 957 LeuThrAlaAlaCys-----CysGluSerLeuSerCysVal 968
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3170 AG-ACAGCACTGTGCTAGTCAGCAGGGTAGCAGGATCTGTGCAGCGCTGCACGAAGGTG 3228
QY 969 IleSerArgSerArgHisLeuLysSer 977
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3229 CCTGTCCAGGAGCCACACCTCCACAGT 3255

RESULT 15
US-09-099-041A-7
; Sequence 7, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (245) ... (3103)
US-09-099-041A-7

Alignment Scores:
Pred. No.: 6,63e-27 Length: 3382
Score: 359.50 Matches: 225
Percent Similarity: 38.30% Conservative: 135
Best Local Similarity: 23.94% Mismatches: 367
Query Match: 5.08% Indels: 214
DB: 4 Gaps: 41

US-10-066-521-6 (1-1344) x US-09-099-041A-7 (1-3382)
```

Qy 189 ThrValValLeuHisGlySerGlyTleGlyTysSerAlaLeuAlaArgArgIleVal 208
||| : : : : :
Db 833 ACCATCTTCATCTGGTGTATCTGGGTGGCAAGTCCATGCTGCTACAGCGGTGCAG 892
Qy 209 LeuCysTrpAlaGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuPro 228
||| : : : : :
Db 893 AGCCTCTGGGCACCGGCGGTAGACGACGAGGTC---AAATCTCTTCCACTTTCGC 949
Qy 229 ValArgGluMetGlnArgLysGluSerSer-----ValThrGluPheIleSer 245
||| : : : : :
Db 950 TCCCGCATGTTTCAGTCTTCAAGGAAGTACAGGCTGTGTCTGACAGACTGCTCTTC 1009
Qy 246 ArgGluTrp-----ProAspSerGlnAlaProValThrGluIleMetSerArg 261
||| : : : : :
Db 1010 AAGCACTACTGTACCCAGAGCGGACCCGAGGAGGTGTTGCTTCTGCTGGCTTC 1069
Qy 262 ProGluArgLeuPheIleAspGlyPheAspAspLeuGlySerValLeuAsnAsn 281
||| : : : : :
Db 1070 CCCCACGTGGCCCTCTTCACTTCGATGCTGGACGAGCTGCACCTGGACTTG----- 1123
Qy 282 AspThrLysLeuCysLysAsp-----TrpAlaGluLysGlnProProPheThr 297
||| : : : : :
Db 1124 GACCTGAGCGCGTGCCTGACAGCTCTGCCCCCTGG---GAGCCTGCCACCCCTGCTC 1180
Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
||| : : : : :
Db 1181 TTGCTGGCCAACTCTCAGTGGGAAGTGTCTCAAGGGGGCTAGCAAGCTGCTCACGCC 1240
Qy 318 Arg-----AspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
||| : : : : :
Db 1241 CGCACAGGCATCGAGGTCCGCGCCACTCTCTCGGAGAAGAGTG----- 1285
Qy 335 LeuLeuValArgGlyLysSerGlyGluGlnArgIleHisLeu----- 348
||| : : : : :
Db 1286 ---CTTCTCGGGGCTTCTCCCCCAGC-----CACCTGCGCGCTATGCCAGGAGG 1333
Qy 349 ---LeuLeuGluArgGlyTleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMet 367
||| : : : : :
Db 1334 ATGTTCCCGAGCGGCGCTGAGGACCGCGCTGTGAGCCAG-----CTGGAG 1381
Qy 368 AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal 387
||| : : : : :
Db 1382 GCCAACCCCAACCTCTGCAGCCTGTGCTGTGCCC-----CTCTTCTGCTGG 1429
Qy 388 AlaLeu-----GlnLeuGln 392
||| : : : : :
Db 1430 ATCATCTTCGGGTGTTCCAGCATTCCGCTGCTGCTTGAAGGCTCACAGCTGCC 1489
Qy 393 AspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAla 412
||| : : : : :
Db 1490 GACTGC-----ACGATGACCTGCACAGATGCTTCTCTCTG 1525
Qy 413 PheVal-----PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsn 429
||| : : : : :
Db 1526 GTCACTGAGTCCATCTGAACAGGATGCAGCCGAGCGCTGGTGGCGG-----AAC 1579
Qy 430 LeuGluGlu-----ArgValValLeuLysArgPheCysArg 441
||| : : : : :
Db 1580 ACACGAGCCCATGTGAGAGACCTCCACGCGCGCGGACACTCTGTGCTGGCGGAG 1639
Qy 442 MetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeuMetVal 461
||| : : : : :
Db 1640 GTGGCCACCGGGCATGTGAGAGAGCCCTCTTGTCTTCCACCCAGGAGGTGAGGCC 1699
Qy 462 GlnGlyLeuGlyGluSerGlu-----LeuArgAlaLeuPheHisMetAsn 476
||| : : : : :
Db 1700 TCCGGGCTCAGAGAGACATGCAGCTGGGCTTCTGCGGGCTTTCGCGAGCTGGGC 1759
Qy 477 IleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisSerLeuGln 496
||| : : : : :
Db 1760 -----CCGCGGGGTGAC---CAGCAGCTCTATGAGTTTTCACCTCACCTCCAG 1807

Qy 497 AspPheCysAlaAlaLeuTyrTyrValLeuGlu----- 507
||| : : : : :
Db 1808 GCCTTCTTTACAGCCTTCTCTCTGCTGACACAGAGGTGGGCACTCAGAGCTGCTC 1867
Qy 508 -----GlyLeuGluIle-GluProAlaLeu----- 515
||| : : : : :
Db 1868 AGTTCTTCCAGGAGTGATGCCCTTGGCGGGGACGACGACCTGCTCTATCTCTCC 1927
Qy 516 -----CysProLeuTyrValGluLysThrLysArgSerMetGluLe 529
||| : : : : :
Db 1928 TTCTCTCCGTTCCAGTCCCTGCGGCGAGTGTCTCGCGGGGAAGACCTCTTCAAGAAC 1987
Qy 529 uLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuVal 549
||| : : : : :
Db 1988 -AAGGAT-----CACTTCCAGTTTCAACAACCTCTCTCTGCGGGCTGTT 2031
Qy 549 lSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVal 569
||| : : : : :
Db 2032 GTCCAAAGCCAAACAGAAACTCTCTGCGGCATCTGTGCCCCGCG---GCAGCCCTGAGAG 2088
Qy 569 lLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsn-----AlaThr 587
||| : : : : :
Db 2089 AAGCGCAAGGCCCTG---TGGGCACACCTGTGTTTCCAGCTGCGGGCTACCTGAAGAG 2145
Qy 587 rThrProGlyAspThrLeuAspAlaPheHis----- 597
||| : : : : :
Db 2146 CTTGCCCGGCTTCAAGTTCGAAAGCTTCAACAGGTGCAGGCCATGCCACGTTTCATCTG 2205
Qy 598 -----CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSe 614
||| : : : : :
Db 2206 GATCTCGCTGTCATCTACGAGACACAGCAGCAGAGGTGGGCGAGCTGGCGCCAGGG 2265
Qy 614 rPheGlnGluValTrpLeuProIleAsnGln-----AsnLeuAspLeuIleAl 630
||| : : : : :
Db 2266 CATCTGCGCCAACTACTCTCAAGCTGACTGCAACGCTGCTCGGCCCACTGCAGCGC 2325
Qy 630 aSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysG 650
||| : : : : :
Db 2326 CTTCTCTCTGCTGTCATCTCTTCCC-----AAGCGGTGGCCCTAGACCTAGACAA 2379
Qy 650 yIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMetArgAs 670
||| : : : : :
Db 2380 CAACAATCTCAACACTACGCGGTGCGG-----GAGCTGAGCCCTGCTTTCAGCCGC- 2431
Qy 670 pLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHi 690
||| : : : : :
Db 2431 ----- 2431
Qy 690 sLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCy 710
||| : : : : :
Db 2432 -CTCACTGTTCTCAGACTCAGCGTAAACACAGATCACTGACGCTGGGTAAAGTGTCTAAG 2490
Qy 710 sAlaLysLeuArgHisProThrCysLysIleGlnThr---LeuMetPheArgAsnAlaG 729
||| : : : : :
Db 2491 CGAAGAGCTG-----ACCAAAATCAAAATTTGTACCTATTGTTGGTTTATACAACACCA 2544
Qy 729 nIleThrPro---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArg 748
||| : : : : :
Db 2545 GATCACCGATGTCGAGCCAGGTAGCTACCAAAATCTGGATGAATGCAAGGCTCTCAC 2604
Qy 748 gSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAl 768
||| : : : : :
Db 2605 GCATCTTAAACTGGGAAAAAACAATAAAGTGAAGGAGGAGATATCTCGCCCTGGC 2664
Qy 768 aLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHi 788
||| : : : : :
Db 2665 TGTGAAGAACAGCAAA----- 2680
Qy 788 sAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSe 808
||| : : : : :
Db 2681 -----TCAATCTCTGAGGTGG 2697
Qy 808 rLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgVal 828

```

Db      :: |||||::||| ||::||| ::::||| |||||
2698  GATGTGGGCAATCAAGTTGGGATGAAGAGCAAAAGCCCTTCGACAGGCTCTGGCG-- 2755

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
828  lserGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCy 848

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
2756  -AACACCCCACTTGACCCCTGAGCTTTGCGTCCAAACGGCATCTCCACAGAAGGAGG 2814

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
848  sGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAs 868

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
2815  AAAGAGCCTTCGAGGGCCCTGCAGCAGAACAGTCTCTAGAAATACTGTGGCTGACCCA 2874

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
868  nAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCy 888

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
2875  AAATGAACCTCAACGATGAAGTGGCAGAGAGTTTGCAGAAATGTTGAAAGTC---AACCA 2931

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
888  sSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLe 908

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
2932  GACGTTAAAGCATTTATGGCTTATCCAGAAATCAGATCACAGCTAAGGGGACTGCCAGCT 2991

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
908  uAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProva 928

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
2992  GGCAGATCGGTTACAGCAACACTGGCATTAACAGAGATTTCCTTAATGGAACCTGAT 3051

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
928  lGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGl 948

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
3052  AAAACACAGAGGAGCCAAAGTCTAT---GAAGATCAGAGCGGATATCTGTTTCTGAGA 3108

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
948  nAsp-----LeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSe 964

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
3109  GGATGCTTTTCCTGTTTCATGGGGTTTTTGCCCTGGAGCCTCAGACGAAATGCCACTCTGG 3168

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
964  rLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAl 984

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
3169  GCAGTCTTTGTGTCAGTGTCTTAAAGGGGCTCGCGAGGGGGACTATCAGGAGTCCAC 3228

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
984  aLeuGlyAspGlyGlyValAlaAlaLeuCys-GluGlyLeuLysGlnLysAsnSerValL 1004

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
3229  TGCCTTCATGATGCAAGCCAGCTTCCTGTGCAGAAAGGTCTGTGTCGGCAA-----C 3279

Qy      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
1004  euThrArgLeuGlyLeuLysAlaCysGly-----LeuThrSerAspCysCys 1019

Db      :: |||||::||| |||||::||| |||||::||| |||||::||| |||||::|||
3280  TCCCTAAGTACCCGCTACAATTCTGCAGAAAAAGAAATGTGTGTTGCGAGCTGTTGT 3335

```

Search completed: July 16, 2004, 06:26:42
Job time : 280 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 15:25:53 ; Search time 7396 Seconds
(without alignments)
5426.551 Million cell updates/sec

Title: US-10-066-521-6;
Perfect score: 7074
Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10066521/runat_13072004_122212_11091/app_query.fasta_1.1543
-DB=Est -QFMT=fastap -SURFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -List=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521 @CGN 1.1 5759 @runat_13072004_122212_11091 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estom:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rpd:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*

29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2653	37.5	3475	11	AK087774 Mus muscu
2	1422	20.1	3359	11	BC021272 Homo sapi
3	1422	20.1	3360	11	BC012789 Homo sapi
4	1217	17.2	791	14	CK000513 AGENCOURT
5	1196.5	16.9	3577	11	AK087843 Mus muscu
6	1182.5	16.7	3218	11	AK054378 Mus muscu
7	1043.5	14.8	1913	11	AK016782 Mus muscu
8	1008	14.2	3793	11	BC036506 Homo sapi
9	962	13.6	3275	11	AK054426 Mus muscu
10	940	13.3	3400	11	AK054264 Mus muscu
11	687	9.7	1119	12	BM454123 AGENCOURT
12	669.5	9.5	814	14	CB228956 AGENCOURT
13	669.5	9.5	2730	11	AK039531 Mus muscu
14	643.5	9.1	1371	29	AY407369 Mus muscu
15	636	9.0	671	10	BB555431 BB555431
16	628	8.9	684	13	BU630481 UI-H-FL0-
17	626	8.8	693	13	BU634350 UI-H-FL1-
18	613.5	8.7	2218	11	AF054176 Homo sapi
19	608.5	8.6	1371	29	AY407367 Homo sapi
20	586.5	8.3	1922	11	AK049352 Mus muscu
21	559	7.9	579	12	BM228619 K0264H01-
22	558.5	7.9	1313	14	CF110534 Shultromi
23	553	7.8	666	9	AV367637 AV367637
24	548	7.7	642	13	BU618831 UI-H-PH1-
25	543	7.7	565	12	BM226108 K0222C06-
26	534	7.5	706	10	BB624558 BB624558
27	517.5	7.3	1434	11	AK014932 Mus muscu
28	507.5	7.2	1160	9	AF522024 AF522024
29	486.5	6.9	589	14	CA559886 K0264H01-
30	485	6.9	511	12	BM230035 K0288H05-
31	485	6.9	513	12	BM227869 K0250F09-
32	483	6.8	519	10	BG071729 H3102B07-
33	483	6.8	519	12	BG084574 H3102B07-
34	482	6.8	515	12	BM227562 K0245F10-
35	481	6.8	512	12	BM226240 K0225A02-
36	481	6.8	901	14	CD515309 AGENCOURT
37	475.5	6.7	583	14	CA559979 K0266B01-
38	473.5	6.7	970	12	BG696072 602658033
39	473	6.7	950	13	BQ951319 AGENCOURT
40	470	6.6	898	13	BU113197 603117379
41	469.5	6.6	593	14	CA558845 K0250F09-
42	459	6.5	1246	29	AY407368 Pan trogl
43	454.5	6.4	643	14	CF913803 B0954D05-
44	449	6.3	1003	12	BM558396 AGENCOURT
45	447.5	6.3	636	14	CF914959 B0971H02-

ALIGNMENTS

RESULT 1
AK087774 3475 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
DEFINITION full-length enriched library, clone:E330019F16 product:mouse WATER
protein (maternal-antigen-that-embryos-require) protein, full
insert sequence.
ACCESSION AK087774
VERSION AK087774.1 GI:26104500
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db	1802	CAATGGACACCCCTGGAGCTCTCTATTGCTATTATTGAGTCTCAGAGTGAAGAGTTCTTGTG	1861
Qy	609	rgLeuAlaLeuAasnSerPheGlnGluValTrpLeuProIleAasnGlnAasnLeuAspLeuI	629
Db	1862	CGCGGGCTCTCAAAACGCTTCCAAAGAGTGTGGCTGCTGATTAAACCAGAAGATGGACCTTGA	1921
Qy	629	leAlaSerSerPheCysLeuGlnHisCysProTrpYLeuArgLysIleArgValaspValL	649
Db	1922	AGGTCTCTTCTACTGTCTCAAGCACCTGTGAGAACTTTGAAGGCAATCCGGGTGGATATCA	1981
Qy	649	ysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMetA	669
Db	1982	GAGACCTCTCTCGTGTGATAAATACTCTCAGAGCTGTGCCCTGTGTGTACT-----GTCC	2035
Qy	669	rgAsp-----LysThrLeuIleGluGluInTrpGluAaspPheCysSerMetLeuG	686
Db	2036	GGGAGACACAATGTAAGCCCTCTCATGGAGTGGTGGGGAACCTCTGCTCTGTGCTTG	2095
Qy	686	lyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaM	706
Db	2096	GCAGCTCCGGAACCTTAGAGGAGCTGACCTTGGCGCAGACATCCTGAGTCAACGGGCCA	2155
Qy	706	etiLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheA	726
Db	2156	TGAAGATACTGTGCTCGAGCTCGGAACTCAGTCTCGAGAAATACAGAACTGACGTTTA	2215
Qy	726	rgAenAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAasnArgA	746
Db	2216	AGAGTGCAGAGGTAGTGTCTGGCTGAAACATCTCTGGAAGCTCCTTTTAGCAATCAA	2275
Qy	746	snLeuArgSerLeuAasnLeuGlyGlyThrHisLeuLysGluIysAspValArgMetAlaC	766
Db	2276	ACTTAAAGTACCTCAATCTAGGGAACACTCCCATGAGGATGATGATCAATGAAGTTAGCT	2335
Qy	766	ysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyL	786
Db	2336	GCGAAGCGCTGAAACATCCAAAGTCTCCCTGGAGACTCTGAGGTTGGATTCTGTGAGT	2395
Qy	786	euThrHisAlaCysTrpLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysS	806
Db	2396	TAACCATCATGTTGTTATGAGATGATCTCCAGCTTCTTATTTCACACAGGCTTAAGT	2455
Qy	806	erLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaL	826
Db	2456	GTCTCAGCTGGCCAAAATAGAGTGGGATGAGTAAAGACATGATATCCCTTGGGAATGCCT	2515
Qy	826	euArgValSerGlnCysAlaLeuGlnLysIleLeuLeuGluAaspCysGlyIleThrAlaT	846
Db	2516	TGAGTAGCTCAATGTGTCTACTGCAAAAGTTGATACTGGACAATCTGGGCTTCACACTG	2575
Qy	846	hrGlyCysGlnSerLeuAlaSerAlaLeuValSerAenArgSerLeuThrHisLeuCysL	866
Db	2576	CCAGCTGCCACTTCTGGTCTCAGCCCTTTTCAGCAACACAGACTTGACACACTGTGCC	2635
Qy	866	euSerAenAasnSerLeuGlyAsnGluGlyValAasnLeuLysCysArgSerMetArgLeuP	886
Db	2636	TGTCAAACAACAGCTGGGAGCTGAAGAGTGCACACAGCTGTGTCACTTCTCGAGGAATC	2695
Qy	886	roHisCysSerLeuGlnArgLeuMetLeuAenGlnCysHisLeuAaspThrAlaGlyCysG	906
Db	2696	CAGAATGTGCTCTCCAGCGGCTGATATGAATCACTGCAACACTGTAGATGATGCTTTATG	2755
Qy	906	lyPheLeuAlaLeuAlaLeuMetGlyAasnSerTrpLeuThrHisLeuSerLeuSerMetA	926
Db	2756	GCCTTCGGCAATGAGACTTGGCAAAACAACAAGACTGACCCACTGAGCTGACCAAGA	2815
Qy	926	snProValGluAAspAasnGlyValLysLeuLeuCysGluValMetArgGluProSerCysH	946
Db	2816	ACCCCTGAGGATGGTGAATGAGCTACTGTGTGAAGCTTTTAAAGGAACCTACTTGT	2875
Qy	946	isLeuGlnAaspLeuGluLeuValLysCysHisIleThrAlaLalaCysCysGluSerLeuS	966
Db	2876	ACCTTCAAGAACTGGAATGAGTGCCTGCAACTCACAAGAACTGCTGGAGAGCTGCTG	2935

Db 389 -----TTATGGTCCAGCAAGTCTGTCACTGA 414
Qy 148 tThrLysPheAlaGluGluAspValArgSerPheGluAenThrAlaAlaAspTr 168
Db 415 GATTCACCTATATCTTTGAGGAGAAAGTCAAGCAAGAA----- 451
Qy 168 pProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProAr 188
Db 452 ----GAATGTGACCATTTGGACCGCTTTTGTCTCCCAAGGAAACTGGGAAACAGCCACG 507
Qy 188 gThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaAArgIleVa 208
Db 508 TACAGTGATATTCAAGGACCAACAGGAATTTGAAAAACGACACTCTCTGATGAAGCTGAT 567
Qy 208 lleuCysTrpAlaGlnGlyGlyLeuTyGlnGlyMetPheSerTyValPhePheLeuPr 228
Db 568 GATGCGCTGGTCGGACAAACAGATCTTTCGGGATAGTTTCTGTACACGTTCTATTTCTG 627
Qy 228 oValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArgGluTr 248
Db 628 CTGCAGAGAACTGAGGGAGTTGCCGCCAACGAGTTTGGTGACTGATTTCCAGAGAGTG 687
Qy 248 pProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIl 268
Db 688 GCCTGACCCCGCTGCTCTCTATACAGAGATGTGTCTCAACCGGAGAGACTTGTGTCGT 747
Qy 268 eIleAspGlyPheAspAspLeuGlySerValLeuAenAen---AspThrLysLeuCyLys 287
Db 748 CATCGACAGCTTCGAAGAGCTGCAGCGCGGCTTGAACGAACCCGATTCCGATCTGTGTGG 807
Qy 287 shApTrpAlaGlyLysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLe 307
Db 808 TGACTTGTATGGAGAAACCGCGGTGCAGGTGCTTCTGAGCAGTTTGTCTGAGGAAGAAT 867
Qy 307 uLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSe 327
Db 868 GTCCCGGAGGCTCCCTGCTCATCGCCATCAAAACCGTGTGCCGAGAGAGCTCCGGGA 927
Qy 327 rGluValValSerProArgTyLeuLeuValArgGlyIleSerGlyGluGlnArgIleHi 347
Db 928 TCAGGTGACGATCTCAGAAATCTACAGCCCGGGGATTCAACGAGAGATGATAGTGTAGT 987
Qy 347 sLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMe 367
Db 988 GTATTCTGCTGTTCTTCAAGACCCGAAAGACCCATGGAAGCTTCAATCTTTGTAAG 1047
Qy 367 tAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCyEva 387
Db 1048 AGAAAGTGAAAGCTGTTTTCATATGCAATCCCGTCTCTGCTGGATCCCTGTGTGAC 1107
Qy 387 lAlaLeuGlnLeuGlnAspValValGlyLysSerValAlaProPheAenGlnThrLeuTh 407
Db 1108 CAGTCTGAAGCAAGAGATGCAGAAAGGAAAGACCTGGCCCTGACCTCCAGAGACTAC 1167
Qy 407 rGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgCy 427
Db 1168 CTCTGTGACTCTCTCTGCTTTAACTGCTTCACTGAGGTGCGGAGGCGCCGAC 1227
Qy 427 sLeuAenLeuGluArgValValLeuLysArgPheCysArgMetAlaValGluGlyVa 447
Db 1228 TCCGCAAAACCCAGCACCAAG-----CTGAAGGCCCTGTGCTCCCTGCGAGAGGGTAT 1281
Qy 447 lTrpAenArgLysSerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluse 467
Db 1282 GTGGACAGACACATTTGAGTTTGTGAAGACGACCTCCGGAGAAATGGGGTGTGTGACGC 1341
Qy 467 rGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTy 487
Db 1342 TGACATCCCTGGCTGCTGGGACCAAGATACTTCTGAAGTACGGGGAGCGGTGAGAGCTC 1401
Qy 487 rTyThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyTyValLeuGl 507
Db 1402 CTACGTGTTCTCCACGTTGTATCCAGGAGTTCTGTGCCGCTTGTCTATTTGCTCAA 1461

Qy 507 uGlyLeuGluIleGluPro-----AlaLeuCySProLeuTy 519
Db 1462 GAGCCACCTTGATCATCTCACCAGCTGTGAGATGTGTACAGGAATTGTAGTTGCCAA 1521
Qy 519 rValGluTyThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLe 539
Db 1522 TTTTGAAGAAAGCAGGAGAGCA-----CATTGAT 1551
Qy 539 uTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVa 559
Db 1552 TTTTGGGGTGTCTTCTAAGTGGCTTTTAAATAAAGAAAGCAAGAAATCTGGATGC 1611
Qy 559 lLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTrpValSerLe 579
Db 1612 GTTTTGTGGCTTCCAACTGTCCCAAGAGATAAAGCAGCAAAATTCACCAAGCTCTGAAGAG 1671
Qy 579 uLeuGlyGlnGlnProAenAlaThrThrProGlyAspThrLeuAspAlaPheHisCyLe 599
Db 1672 CTTAGGGGAGCGTGGCAATCTCAGGAGCAGGTGATTCCTTGGCGATATTTTACTGTCT 1731
Qy 599 uPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAenSerPheGlnGluValTr 619
Db 1732 CTTTGAATGCAGATCTGCTTGTGAAGCAGCAGTGAACCTCTCCAGAGAGCTAA 1791
Qy 619 pLeuProIleAenGlnAenLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPr 639
Db 1792 CTTTCATATTATGACAACGTGGACTTGGTGTCTCTGCTACTGCTTAAATAATGCTC 1851
Qy 639 oTyLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGl 659
Db 1852 CAGCTTCAGGAACCTCTGTTTTCCGTTCAAAATGTCTTAAAGAAAGAGGATGAACACAG 1911
Qy 659 uAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGl 679
Db 1912 CTCT-----ACGTGGATTACGCTCATC-----TGTGGCA 1944
Qy 679 uAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSe 699
Db 1945 TCACATCTGCTGTCTCACCAGCGGGCACCCTCAGAGAGCTCCAGGTGTCAGGACAG 2004
Qy 699 rIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLy 719
Db 2005 CACCTTCAGGAGTCGACCTTTGTGACCTGCTGTGAACCTGAGGATCCAGCTCTCG 2064
Qy 719 sIleGlnThrLeuMetPheArgAenAlaGlnIleThrProGlyValGlnHisLeuTrpAr 739
Db 2065 CTTTCAGAGCTTGAATAATAACGTTTCTTCTGGCCAGAGTGTCTGCTCTTTGA 2124
Qy 739 gIleValMetAlaAenArgAenLeuArgSerLeuAenLeuGlyGlyThrHisLysLysGl 759
Db 2125 GGTGCTCTTTTATCAGCGACACTTGAATAACCTGAGCTTACCCTCAGAAACTCTCTCG 2184
Qy 759 uGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLe 779
Db 2185 TGATGATCATGCTCCCTCTGTGATGCTTGAATCCCA----- 2224
Qy 779 uArgLeuAspCysCysGlyLeuThrHisAlaCysTyLeuLysIleSerGlnIleLeuTh 799
Db 2224 ----- 2224
Qy 799 rThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVa 819
Db 2225 -----GCAGGCAAC----- 2233
Qy 819 lMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGl 839
Db 2234 -----GTCAAAGAGCTAGCGCTGT 2253
Qy 839 uAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAenAr 859
Db 2254 AATTGTACACTCTCACCCCATTTGATTGTGAAGTCTTGTGCGCTTCTTAACCAACAAACA 2313

```
Qy 859 gserLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValaenLeuLe 879
Db 2314 GAAGCTGACGATCTGATGATCTCTGCAACAGTTA---GACACAGCGGTGCCCTTTT 2370
Qy 879 uCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAenGlnCysHi 899
Db 2371 GTGTGAAGCCCTGTGCACCCAGACACGCTCTGTGTATACCTGATGTTGGCTTTCGCA 2430
Qy 899 sLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuTh 919
Db 2431 CCTCAGCAGCAGCTGCTCGCAATACATCTCTGAATGCTTCGTGTAACAGACGCGTGC 2490
Qy 919 rHisLeuSerLeuSerMetAsnProValGluAspAsnGlyVallylsLeuLeuCysGluVa 939
Db 2491 CTATCTAGACCTCAGTGCATGCTCCTGAAGACGAGGAGTGAAGAACTCTCTCGGAGGC 2550
Qy 939 lMetArgGluProSerCysHisLeuGlnAspLeuGluLeuVallylsCysHisLeuThrAl 959
Db 2551 CTTGAAACATCCGACTGCTCGCTGGATTCACTGTGTTGGTAAATGTTTATCACTGTC 2610
Qy 959 aAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAs 979
Db 2611 TGCTGGCTGTGAAGACCTCTGCTCTGCTCATCAGCAATCAAAACCTGAAGATTCTGCA 2670
Qy 979 pLeuThrAspAenAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGl 999
Db 2671 AATGGGTGCATGAAATCGAGATGCGGTGTCAGCTGTGTGTCGGCTCTGACGCA 2730
Qy 999 nLyAsnSerValLeuThrArgLeuGlyLeuAlaCysGlyLeuThrSerAspCysCy 1019
Db 2731 TACGATTGCCCTTAGAGATTCTGGGTGGAAGATTGGGTAAACGAGCACCTGCTG 2790
Qy 1019 sGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAenLeuValGl 1039
Db 2791 TAAGGATCTCGCGTCTGTCTTCACCTGCAGTANGACCTCGACGAGCTCAACCTGACCTT 2850
Qy 1039 nAsnAsnPheSerProLyGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSe 1059
Db 2851 GAACACCTTGACACACACAGGGGTGTGTACTCTGTGAGCCCTGAGACACCCAGAGTG 2910
Qy 1059 rAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLe 1079
Db 2911 TGCCCTGCGAGTGTGCGGCTGAGAAAACCTGATTTGATGAGGAACCCAGGACACTCT 2970
Qy 1079 uGluGluValGlnLeuLeuLysProArgValValIleAsp----- 1092
Db 2971 GACGGCTGAGGAGAGAGAAATCCTAACCTGACATCAGATGACTGTGACACAATCAC 3030
Qy 1093 -----GlySerTrpHisSerPheAspGluAspArgHisLy 1105
Db 3031 AAGGTAGAGATCTGATTCGAGGAACCTGG----- 3061
Qy 1105 sIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCysAlaLeuLeuTrpGl 1125
Db 3062 ---GCTCTGACTCGAACACCTGCAAGAGGACGAGCTGGGAC---CGTTACTTTACATGA 3114
Qy 1125 yMetAsnProGluGlnIlyLysArgValSerLeu 1136
Db 3115 CACTGCAACCCAGGAGATACAAATCATTTGACATC 3148

RESULT 3
LOCUS BC012789 3360 bp mRNA linear HTC 17-DEC-2003
DEFINITION Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift errors.
ACCESSION BC012789
VERSION BC012789.2 GI:33872518
KEYWORDS HTC
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3360)
```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zdobych, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3360)

Strausberg, R.

Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:15215377.

Contact: MGC help desk

Email: cgapsb-remail.nih.gov

Tissue Procurement: ARCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

http://www.systemsbio.org

contact: anadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettenan, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 25 Row: k Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 19745161

This clone has the following problem: frame shifted.

Location/Qualifiers

1..3360

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3957172"

/tissue_type="placenta, choriocarcinoma"

/clone_lib="NIH MGC 21"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

ORIGIN

Alignment Scores:

Pred. No.: 3 35e-133 Length: 3360

Score: 1422.00 Matches: 349

Percent Similarity: 48.00% Conservative: 204

Best Local Similarity: 30.30% Mismatches: 448

Query Match: 20.10% Indels: 151

DB: 11 Gaps: 15

US-10-066-521-6 (1-1344) x BC012789 (1-3360)

Qy 9 PheSerSerTrpGlyLeuGlnTrpCysLeuTyrgluLeuAspLysGluGluPheGlnThr 28

Db 76 TTCTCTGATTTTGGTCTTATGCTGATCTGAGGAGCTCAAAAAGGAGGAGTTTCAGGAAA 135


```
Db 2125 GGTGCTCTTTTATCAGCCAGACTTGAATACCTGAGCTTACCTCCCTCAGAAACTCTCTCG 2184
Qy 759 uGluaspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLe 779
Db 2185 TGATGACATCATCAGGTCCCTCTGTGATCCCTTGAACATACCCA----- 2224
Qy 779 uArgLeuaspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2224 ----- 2224
Qy 799 rThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVa 819
Db 2225 -----GCAGGCAC----- 2233
Qy 819 lMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGl 839
Db 2234 -----GTCAAAGAGCTAGCGCTGGT 2253
Qy 839 uAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnAr 859
Db 2254 AAATTGTCACCTCTCACCCATTGATGTGAAGTCTTGTCTGGCCTTCTTAACCAACAA 2313
Qy 859 gSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLe 879
Db 2314 GAAGTGAAGTATCTGAATGATCTCTGCAACAGTTA---GACACAGCGTGGCCCTTTT 2370
Qy 879 uCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHi 899
Db 2371 GTGTGAAGCCTGTGCAGCCAGACAGCGTCTGTGTATCTGATGTTGGCTTTCTGCCA 2430
Qy 899 sLeuAspThrAlaGlyCysGlyPheLeuAlaLeuMetLeuValAsnSerTrpLeuThr 919
Db 2431 CCTCAGCAGCAGTCTCGAATACATCTCTGAATGCTTCTGGTAAACAGACGCTGGC 2490
Qy 919 rHisLeuSerLeuSerMetAnProValGluAspAsnGlyValLysLeuLeuCysGluVa 939
Db 2491 CTATCTAGACCTCAGTGCCATGCTCTGAGAGCAGAGGACTGAACCTCTCTCGAGGC 2550
Qy 939 lMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAl 959
Db 2551 CTTGAAACATCCGACTCTCGCTGGATTCTCTGTTTGGTAAATGTTTATCACTGC 2610
Qy 959 aAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAs 979
Db 2611 TGCTGGCTGTGAAGACCTCGCCTCTGCTCTCATCAGCAATCAAAACCTGAAGATCTGCA 2670
Qy 979 pLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGl 999
Db 2671 AATTGGGTGCATGAAATCGAGATGCGGTGTGCAGCTGTGTGTCGGGCTCTGACGCA 2730
Qy 999 nLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCy 1019
Db 2731 TACGGATTGCGCCTTAGAGATTCTTGGGTGGTGAAGAATGTGGGTAAACGACGACCTGCTG 2790
Qy 1019 sGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGl 1039
Db 2791 TAAGGATCTCCGCTGTGTTCTCAGCTGCAATGAAACCTCGAGCAGCTCAACCTGACCTT 2850
Qy 1039 nAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSe 1059
Db 2851 GAACACCTTGGACACACAGGGGTGGTGTACTCTGTGAGGCCCTGAGACACCCAGAGTG 2910
Qy 1059 rAsnLeuGlnIlelleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLe 1079
Db 2911 TGCCCTGAGGTGTCTCGGGCTGAGAAAACCTGATTTTATGAGGAAACCCAGGCACTTCT 2970
Qy 1079 uGluGluValGlnLeuLysProArgValValIleasp----- 1092
Db 2971 GACGGCTGAGGAGAGAGAAATCTTAACCTGACCATACAGATGCTGTGACACAACTAC 3030
Qy 1093 -----GlySerTrpHisSerPheAspGluAspArgHisLy 1105
```

```
Db 3031 AAGGGTAGATCTGATTGCGAGGAACCTGG----- 3061
Qy 1105 sIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCysAlaLeuLeuTrpGl 1125
Db 3062 ---GCTCTGACTCGAACACCTGCAAGGACAGGACTGGGAC---CGTTACTTATCATGA 3114
Qy 1125 yMetAsnProGluGlnLysLysArgValSerLeu 1136
Db 3115 CACTGCACCCAGAGATACAAATCATTTGACATC 3148

RESULT 4
CK000513 791 bp mRNA linear EST 26-NOV-2003
LOCUS AGENCOURT_16368905 NIH_MGC_221 Homo sapiens cDNA clone
DEFINITION IMAGE:30708637 5', mRNA sequence.
ACCESSION CK000513
VERSION CK000513.1 GI:38526547
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 791)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAMI075 row: c column: 14
High quality sequence stop: 688.

FEATURES
source
1..791
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708637"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5kb. Adaptors 5' (AATTCGACGACGAGG)3' and 5'd
(CCTCGTGGCG)3'. 3' linker sequence - GCGGCGCTGAGAGCC T18.
Sequencing primers 3' end: T3 promoter primer 5'd
(AATTAACCTCTACTAAGGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2..2e-113 Length: 791
Score: 1217.00 Matches: 255
Percent Similarity: 98.08% Conservative: 1
Best Local Similarity: 97.70% Mismatches: 5
Query Match: 17.20% Indels: 5
DB: 14 Gaps: 0
```

```
US-10-066-521-6 (1-1344) x CK000513 (1-791)

QY 190 ValValLeuHisGlyLysSerGlyLeuGlyLysSerAlaLeuAlaArgArgIleValLeu 209
Db 8 GTGGTTCTGCGACGGAAGTCAGGAATTTGGAAATCGGCTCTAGCCAGAGGATCGTGTG 67

QY 210 CysTrpAlaGlnGlyLeuTyrlngLingMetPheSerTyrlValPhePheLeuProVal 229
Db 68 TCTGTGGCCACAGGTGGACTCTACACGAGGAATGTTCTCTACGTCTTCTTCTCCCGCTT 127

QY 230 ArgGluMetGlnArgLysLysGluSerValThrGluPheIleSerArgGluTrpPro 249
Db 128 AAGAGAGATGACGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 187

QY 250 AspSerGlnAlaProValThrGluLeuMetSerArgProGluArgLeuLeuPheIleIle 269
Db 188 GACTCCAGGCTCCGGTGAAGGAGAGATCATGTCGCCAGCAGAAAGGCTGTGTTCATCAT 247

QY 270 AspGlyPheAspAspLeuGlySerValLeuAenAenAspThrLysLeuCysLysAspTrp 289
Db 248 GACGGTTCGATGACCTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGCAAAAGCTGG 307

QY 290 AlaGluLysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuPro 309
Db 308 GCTGAGAGACAGCCCTCCGTTTCACTCATACAGCAGTCTGCTGAGGAAGGCTCTGCTCCCT 367

QY 310 GluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluVal 329
Db 368 GAGTCTTCTCTGATCGTCACTGACAGAGCGTGGGCACAGAGAGAGCTCAAGTCAGAGGTC 427

QY 330 ValSerProArgTyrlLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeu 349
Db 428 GTGTCTCCCGTTACTGTGTAGAGGAATCTCCGGGGAACAAGAATCCACTTGCTCTC 487

QY 350 LeuGluArgGlyLeuGlyHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsn 369
Db 488 CTGTAGCGGGGATTTGGTGGATCATCAGAGACACAAAGGTTGCGTGCATCATGAACAA 547

QY 370 ArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeu 389
Db 548 CGTGAGCTGCTCGACAGTGCAGGTCGCGCGCGTGGGCTCTCTCATCTGGGTGGCCCTG 607

QY 390 GlnLeuGlnAspValValGlyLysSerValAlaProPheAenGlnThrLeuThrGlyLeu 409
Db 608 CAGCTGCGAGAGCTGTGGGGAGAGCGTGGCCCTTCAACAAACGCTCACAGCCCTG 667

QY 410 HisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsn 429
Db 668 CACGCGCGCTTTTGTGTAT-CAGCTCACCCCTCGAGGCGTGGTCCGCGCTGCTCMAAT 726

QY 430 LeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsn 449
Db 727 CTGGAGGA-AGAGTTGTCTGTAG-CGCTTCTGCCCTTG-GCTGTGGA-GGAGTGTGGAA 782

QY 450 Arg 450
Db 783 AAG 785

RESULT 5
AK087843
LOCUS
DEFINITION
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330028A19 product:Similar to
PAN2 protein, full insert sequence.
ACCESSION
AK087843.1 GI:26352787
VERSION
AK087843
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Carninci, P. and Hayashizaki, Y.
```

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takanashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

URL: http://location.gsc.riken.go.jp/

Location/Qualifiers

1. 3577

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:E330028A19"

/db_xref="MG:2426651"

/db_xref="taxon:10090"

FEATURES

Source

JOURNAL
MEDLINE
PUBMED
REFERENCE

AUTHORS

TITLE

JOURNAL
REFERENCE
AUTHORS

TITLE

JOURNAL
REFERENCE
AUTHORS

TITLE

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length CDNA's
Nature 420, 563-573 (2002)
6 (bases 1 to 3218)

Fukuda,S., Aizawa,K., Akamura,T., Arakawa,T., Bono,H., Carninci,P., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi.K., Nomura,K., Numasaki,R., Onno.M., Ohsato.N., Okazaki.Y., Saito.R., Saitoh.H., Sakai.C., Sakai.K., Sakazume.N., Sano.H., Saeki,D., Shibata,K., Shinagawa.A., Shiraki.T., Sobabe,Y., Tagami,M., Tagawa,K., Takahashi.F., Takaku-Akahira.S., Takeda,Y., Tanaka.Y., Tomaru.A., Toyota.T., Yasunishi.A., Muramatsu,M. and Hayashizaki,Y.

Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

CNDA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1..3218
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:E330021B02"
/db_xref="MGI:2426517"
/db_xref="taxon:10090"
/clone="E330021B02"
/sex="female"
/tissue_type="ovary"
/clove_lib="RIKEN full-length enriched mouse cdna library"
/dev_stage="2 days pregnant adult"

misc_feature
1..3218
/note="Similar to PAN2 protein (SPTR|Q99MW0, evidence:
FASTY, 78.9%ID, 100%length, match=2244)"

ORIGIN

Alignment Scores:

Pred. No.:	Length:
Score:	1182.50
Percent Similarity:	49.42%
Best Local Similarity:	28.27%
Query Match:	16.72%
DB:	11
Gaps:	17

US-10-066-521-6 (1-1344) x AK054378 (1-3218)

PhaserSerTyrclyTeuGlntPrCysLeutyrgluLeuAsplysGluGlnThr 28


```

Db      833  ----- 833
Qy      824 AspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIle 843
Db      833  ----- 833
Qy      844 ThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHis 863
Db      834  ----- 834
Qy      864 LeuCysLeuSerAsnAsnSerLeuGlyAsnGlyValAsnLeuLeuCysArgSerMet 883
Db      895 CTGTGCTGTCTAAACACAGCCTGGGAGTGAAGAGTGCACACCTGTGTGCTTCCTG 944
Qy      884 ArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAla 903
Db      945 AGGAATCCAGATGTCTCTCCAGCGCTGATCTGAATCATTTGCAACATTGTAGATGAT 1004
Qy      904 GlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeu 923
Db      1005 GCTTATGGCTTCTGTGCAATTGAGACTTGTCAACACACACAAAGCTGACCCACCTGAGCCTG 1064
Qy      924 SerMetAsnProValGluAsnGlyValLysLeuLeuCysGluValMetArgGluPro 943
Db      1065 ACCATGAACCCGTAGGGAGTGTGCAATGAAGTACTGTGTGAAGCTTTTAAAGGAACCT 1124
Qy      944 SerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGlu 963
Db      1125 ACTTGTTACCTTAAAGAACTGAACCTAGTGGACTGCCAACTTATGCAGAACTGCTGTGAG 1184
Qy      964 SerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsn 983
Db      1185 GACCTGGCTGTATGATCACAAACCAAGCACTTAAAGAGTTTGGATCTTGGTAACAAC 1244
Qy      984 AlaLeuGlyAspGlyValAlaAlaLeuCysGlyLeuGlyLeuLysGlnLysAsnSerVal 1003
Db      1245 GCCCTGGGTGACAAAGGAGTCTAAACCTGTGCGAGGAGCTGAAGCAATGATGAGCTCC 1304
Qy      1004 LeuThrArgLeuGlyLeuAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSer 1023
Db      1305 CTGAGGAGACTTGGTGTGGGGCATGTGATGTGACTTCCATTTGCTGTGATCTATTGTCA 1364
Qy      1024 LeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSer 1043
Db      1365 TTGGCCATCTCTTGAACCCCTCACCTGAACAGCCTTAAACCTGTATGATGATGACTTCAGT 1424
Qy      1044 ProlLysGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIle 1063
Db      1425 ACATCGGGGAAGTTGAAGCTGTGCTCTGCGTTCCAAATGCCCTGTCTTAACCTGGGGATA 1484
Qy      1064 IleGlyLeuTrpLysTrpGlnTrpProVal-----GlnIleArgLysLeuLeu 1079
Db      1485 ATTGGGTC-TGGCACTTGGGCTCACCTGCATTAACTTTCATGATTAAGAGAACTA 1543
Qy      1080 GluGluValGlnLeu-----LeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db      1544 AGGGACAGGATTTAAACAATCTGTCTCCAAATTTAATC-----CATACA 1588
Qy      1098 Phe 1098
Db      1589 TTT 1591

RESULT 8
BC036506
LOCUS      3793 bp      mRNA      linear      HTC 19-NOV-2003
DEFINITION Homo sapiens NACHT, leucine rich repeat and PYD containing 11, mRNA
(cDNA clone IMAGE:5265613), with apparent retained intron.
ACCESSION BC036506
VERSION   BC036506.1 GI:23243476
KEYWORDS  HTC.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE AUTHORS

Mammalia; Butheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 3793)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Sapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carrinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Vallalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahl,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvachenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE PUBMED

22388257
 12477932

REFERENCE AUTHORS

2 (bases 1 to 3793)
 Direct Submission
 Strausberg,R.
 Submitted (09-AUG-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 73 Row: m Column: 9
 This clone has the following problem: retained intron.

FEATURES

source
 Location/Qualifiers
 1..3793
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5265613"
 /tissue_type="Testis"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="vector: pbluescript"

ORIGIN

Alignment Scores:
 Pred. No.: 1.24e-90 Length: 3793
 Score: 1008.00 Matches: 269
 Percent Similarity: 49.47% Conservative: 200
 Best Local Similarity: 28.38% Mismatches: 402
 Query Match: 14.25% Indels: 77
 DB: 11 Gaps: 19
 US-10-066-521-6 (1-1344) x BC036506 (1-3793)
 Qy 177 AlapheaSpSer---AepArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
 |||:::|||||
 :::: :::

Db 787 GCCTATGATTTACACGCTATTATTTCAGCAACAATCTCAATGTGTTCTCTGATGGAGAG 846
Qy 196 SerGlyIleGlyLeuSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 847 AGAGCATCTGGAAGAACTATTGTTATAAATCTGGCTGTGTTGAGGTGGATCAAGGGTGAG 906
Qy 216 LeuTrpGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
Db 907 ATGTGGCAGACATGATCTGTACGTCTGTTACCTCACTTCTCAGCAATTAACACAGATG 966
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 967 ACCAACACAGCAGCTGTGGCTGAGCTAATCGCCAGGACTGGCTGACGGCCAGCTCCCAT 1026
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeu 275
Db 1027 GCAGACATCTGTCTGATCCCAAGAACTCTCTTTCATCTCGAGGACTTGGGCAACATA 1086
Qy 276 GlySerValLeuAsn---AsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 1087 AGATTTCGAGTTAATGTCAATGAAGTCTTTGTGTAGTACACAGCACCAGAAAGTTCCC 1146
Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 1147 ATTCAGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
Qy 315 ValThrValArgAspValGlyThrCluLysLeuLysSerGluValValSerProArgTyr 334
Db 1207 ATCTCTCAAGGCCACACGTGGGGAATAATGTAAGAGCTTCTTGAAGAGGTAGATTGC 1266
Qy 335 LeuLeuValArgGlyIleSerGlyCluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 1267 TGCAGACCTTGCAGCTCTCGAATGGGAAGG---GAGATATATTTAACCTCTTTCTTT 1323
Qy 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeu--- 373
Db 1324 AAAGACCCCGCAGAGGGCTCGCCAGCCCTCCAGCTGTGTACATGAGGATGAATACTCGT 1383
Qy 374 AspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAsp 393
Db 1384 GGTCTGTGCCAGTCGCCATCTTATGTGTGATCAGCTGTACTGTCTGAAAGCGGCAGATG 1443
Qy 394 ValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPhe 413
Db 1444 GACAAGGGGGCTGACTTCCAGCTCTGTGCTGCCAACACCCACTGATCTACATGCCACTTT 1503
Qy 414 ValPheHisGlnLeuThrProArg---GlyValValArgArgCysLeuAsnLeuGluGlu 432
Db 1504 CTGTGCTGATGCTGTGACATCAGAGGCTGGAATCTTACTGCTCAATCAGTATCACTAGGT--- 1560
Qy 433 ArgValValLeuLysAspPheCysArgMetAlaValGluGlyValTrpAsnArgLysSer 452
Db 1561 -----CTCTTAACAGTCTGTGTTGTGCTGCGAGGAGACTGTTCTGAGCACCCTG 1614
Qy 453 ValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeu 472
Db 1615 AATTTCAGTGTGAAGACTTCAGATGTGTTGGTGTACTGAGGCTGATGCTCTGTGTTG 1674
Qy 473 PheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHis 492
Db 1675 CAGCGCCGGAATATCTTTTCCGAGCAACACTCATAAAGACCGTTCACAGTTTCATACAC 1734
Qy 493 LeuSerLeuGlnAspPheCysAlaLeuTyrTyrValLeuGluGlyLeuGluIleGlu 512
Db 1735 TTGAACGTCTCAGGAGTTTGTACAGCCATTCGATTTCTGTG-----ATGGCAGTACCC 1785
Qy 513 ProAlaLeuCysProLeu-----TyrValGluLysThrLysArgSerMetGlu 528
Db 1786 AACTATCTGATCCCTCAGGCAGCAGAGAGTATAAGAGAGAGAGACAATACTCTGTAC 1845
Qy 529 LeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeu 548
Db 1846 TTTAATCAA-----GTGTTTACTTTTCTTTTGTGCTTCTCTGCTGCTCTGCT 1878

Qy 549 ValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGly 568
Db 1879 CTAATGCAACACAGGAAAGATTCTTTGACATCTCTTTGGATACCCAGCTACCGATGGTA 1938
Qy 569 ValLysGlnLysLeuLeuHisTrpValSerLeu-----LeuGlyGlnGln 583
Db 1939 GACAGCTTCAAG-----TGTACTCGTGGGATACATGAACATTTTGGACCGTGAC 1989
Qy 584 ProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGln 603
Db 1990 CCGGAAAGTTGACG-----CACCATATGCTTTGTTTACTGCTCTATGAGATCGG 2043
Qy 604 AspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsn 623
Db 2044 GAGAAGAAATTTGTGAAGACGATTGTGGATGCTCTCATGGAGGTACAGTTTACCTTCAA 2103
Qy 624 GlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLys 643
Db 2104 TCAGACAAGATATGATGGTCTCATTTACTGTTGATTTACTGCTCTACCTGAGGACA 2163
Qy 644 IleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProVal 663
Db 2164 CTTAAGTTGATGCTTCAGCGCATCTTTCAAAACAAAGAG-----CCACTT 2208
Qy 664 Val---ProLeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCys 682
Db 2209 ATAAGGCCAACTGCTAGTCAAAATGAAGAGCTTTGTC-----TACTGGAGAGAGATCTGC 2262
Qy 683 SerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThr 702
Db 2263 TCTCTTTTATACAAATGGAGACCTTCGGGAGCTGCATATCTTTGACATGACCTTAAT 2322
Qy 703 GluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThr 722
Db 2323 GGTATTTCAAGAGGATTTCTGTCTAAAGCCCTCGAGCAITCTAGCTGTAACTTCGCACA 2382
Qy 723 LeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMet 742
Db 2383 CTAAGTTGCTTATGTCGACTGCTTGTGTTTGAAGACTTACTCAAGGCTTTGGCT 2442
Qy 743 AlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspVal 762
Db 2443 CGTAATCGGAGCTGACATACCTGAGTATCAACTGCTACGTCATTTCCCTAAATATGTTT 2502
Qy 763 ArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAsp 782
Db 2503 TCATCTTGTGATGACATCTGTGACAGGCCACATGCCAAATAAGTCATCTGAGAATACAC 2562
Qy 783 CysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerPro 802
Db 2563 -----AGCCCAACGGCAGCGCTCACG----- 2583
Qy 803 SerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeu 822
Db 2584 -----CCTGGAATCCCAACACTT 2601
Qy 823 SerAspAla-LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGln 842
Db 2602 TGGGAGCGCGAGCGCGGCGAGATCACCTGAGCTTTGATGAAA-----TGTGA 2646
Qy 842 YlleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThr 862
Db 2647 TTTGCGAGCCAGCGAATGCGAAGAAATCGCTCTCTCTCATGTCGGGGAGTCTGAG 2706
Qy 862 rHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSe 882
Db 2707 AAAACTGACCTTATCCAGCAATCCGCTGAGGAGCGAGCGGATGACATACATGCTGTGATGC 2766
Qy 882 rMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThr 902
Db 2767 CTTGCTTATCCCACTGACCTTATATACCTGCTGTTAGTCTTCTGCTGCTCTACTGA 2826

```
QY 902 rlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSe 922
Db 2827 AAATTGCTGCGCCCTGGAGAGTGCTTCTGTTTCAGCCCACTTAAGACCACTAGA 2886
QY 922 rLeuSerMetAnProValGluAspAsnGlyVallyLeuLeuCysGluValMetArgG1 942
Db 2887 CTGTGTGGAATCGCTTAAANAATACGGAGTGTTCATGATGAGCTTCCCTGCTGTT 2946
QY 942 uProSerCysHisLeuGlnAspLeuVallyCysCysHisLeuThrAlaAlaCysCy 962
Db 2947 TCCAACTGTCTAGAGGAGCTTCATCTGCTGCTGCTTCTTTAGCAGCGATATCTG 3006
QY 962 sGluSerLeuSerCysValIleSerArgSerArgHisLeuLeuSerLeuAspLeuThrAs 982
Db 3007 TCAATATATATGCTATGCTTACTAATAAGAACTGAGAGCTGAGAGCTGGAGTGGG 3066
QY 982 pAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlyLeuLeuGlnLeuAsnSe 1002
Db 3067 CACAAATAGAGATGCGAGTAATCAGCTGCTATGCTGTTGAGACATCCCACTG 3126
QY 1002 rValLeuThrArgLeuGlyLeuAlaCysGlyLeuThrSerAspCysGluAlaLe 1022
Db 3127 CATGTGTGTAATATGGCTAGAGAGTGCATGTTAACAGTGCCTGCTGCTGCTCT 3186
QY 1022 uSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuLeuValGlnAsnAspH 1042
Db 3187 TGCCTCTGCTTACCAACCAACAACTAGAAAGACTCAACTTGTCTCAAAATCACTT 3246
QY 1042 eSerProlyGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSerAsnLeuG1 1062
Db 3247 GGGCAATGATGAGTGTGCAAACTCTTGGAGAGCTTGATGAGCCAGATGTGTACTTAA 3306
QY 1062 nIleLeuGlyLeuThrPlyTrpGlnTy-ProValGlnIleArgLysLeuGluGluVa 1082
Db 3307 GGTAGTGGCTTCCATTAAGTGGCTGACACACAAACCCAGCAGTGTGAGCAGCTGT 3366
QY 1082 lGlnLeuLeuPysProArgValValIleAspGlySerTrpHisSerPheAspGluAspAs 1102
Db 3367 AAAGGAAGAAACCCAGTGTGATCTTCTGCTGAAACTTGTCTTTAAAGGAAGGC-- 3424
QY 1102 pArgHisIleGlyLeuThr 1109
Db 3425 ----AGAGAAATTTGGTGACA 3442

RESULT 9
AK054426
LOCUS
DEFINITION
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330024M09 product:hypothetical
RNI-like structure containing protein, full insert sequence.
ACCESSION
AK054426
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
```

```

Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, K., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3275)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
source
1. .3275
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E330024M09"
/db_xref="WGI:2426593"
/db_xref="MGI:2426593"
/clone="E330024M09"
/sex="female"
/tissue type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
169. .2760
/notes="unnamed protein product; hypothetical RNI-like
structure containing protein (SCOP|52047, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC35775.1"
/db_xref="GI:26344237"
/translation="MASLFSDFGFIFWYKELNKFMYFKELLHIEHLQMLQKQISWT
EVKEASREDLAILLVKHCQGNQAWDTTFRVQFMIGRNVITNRTATGEIAAHSITYRAHL

```

KEKLTDCSRKFNISQNFQDEYDHELENLVPNGTENNPKWVLQGVAGIGKTLILK
 NMLIVDSGWNKYSFYFCCHDVQKQTAADIILSREWPSPQMBEILSPOE
 KLMFIDISLEGMEWNTQDSQJCYCNHCKQVNVLLSKLIPSSLLSCE
 TFKDQVITNVRITQYFC KENNIMCFHSJLQDRTAQAASJLIRENEQLFVFCQA
 PVQYVWATCLNIEISGKQPSICRR7TSLFYTHLPHNAQNPNSSELDLND
 PFCVABGVMTDLSVNEBAEGLMTDMSDITPLADJLGSLOSRESENSJILPHPSQV
 EFCAMFVHLSHMDHSQGVAFETGFLTPFLANKQVPLGCGCFGLLHETVQAMK
 EAPFGYHLSKELRRQFLMLLELLLOTLPHDPVKKNTMFPYCLFEMEVEFVQAMK
 REQIDVVGKJEDTIVAEVLSHGSALTSLSAQNVNLEEGQXKMLIWHQICSV
 FLRNQDKTLRIEDTIFNEBVFKISYLVKNSCSKTLTVAYNSFLCDKRLFLELT
 SYNLEELRGLSDHVMCLDILNQAQCNIRILDJLANSCLGSHWSDVYK
 KSJRLNINSYNNLQBSGLAMCRLTPNSALSHIRLJGACGCKDLASTFTRVK
 CLRRINLAKNSIGFSGFLVCLCKAMKQDQCTLTYELKLRWADFSDSDQSFELSEMERNKI
 L-LSYENK

polyA signal

polyA site

ORIGIN

Alignment Scores:

Pred. No.:	4.88e-86	Length:	3275
Score:	962.00	Matches:	270
Percent Similarity:	45.56%	Conservative:	181
Best Local Similarity:	27.27%	Mismatches:	416
Query Match:	13.60%	Indels:	123
DB:	11	Gaps:	20

US-10-066-521-6 (1-1344) x AK054426 (1-3275)

9	PheSerSerTyr-GlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThr	28
181	TTTTTCAGATTTTGGCTTTATATGGTACTCGAAAGAGCTTAAACAAGATAGAAATTCATGTAT	240
29	PheLysGluLeuLeuLysLysSerSerGluSerThrThrCysSerLileProGlnPhe	48
241	TTTAAAGGANTTACTCATCATGAAATTCGCAATGGCTTAAAAACAGATTTCTTGGA	300
49	GluLeuGluAsnAlaValGluCysLeuAlaLeuLeuLeuHisGluTyrGlyAla	68
301	GAAGTAAGGAAGCATCTCGGGAAGACCTTGCATCTTACTGGTAAAAACATTTGACGGG	360
69	SerLeuAlaTrpAlaThrSerLilePheGluAsnMetAsnLeuAArgThrLeuSer	88
361	AATCAAGCTTGGGATACGACCTTCCAGATCTTCCAGATGATGGGAGGAATGTATACACT	420
89	GluLysAlaAArgAspMetLysLysLileSerGlnAlaMetGluGlnGluGlyAlaThr	108
421	AATAGGGCA-----	429
109	AlaAlaGluThrGluGluGlnGluLeuSerGlnAlaMetGluGlnGluGlyAlaThrAla	128
430	-----ACAGGAGAGATTGCT	444
129	AlaGluThrGluGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMet	148
445	GCACACTCAACA-----ATATATCGAGCTCATTTAAAG	477
149	ThrLysPheAlaGluGluGluAspValAArgSerPheGluAsnThrAlaAlaAspTrp	168
478	GAGAAGCTGACC-----CATGATTGTCTAGAAAAGTTCACATCAGTATTTCAGAAATTC	531
169	-----ProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheAArg	186
532	TTCCAGGATGAATATGATCATCTTTCGAGAACCTTCTGTGACCAATGGAACTGAAAAACAAT	591
187	ProAArgThrValValLeuHisGlyLysSerGlyTleGlyLysSerAlaLeuAlaAArgAArg	206
592	CCAAAGATGTTGTCTCTCGAAGGTGTAGCTGGANTTGGCAAGACTATTCTGTTAAAAAAT	651
207	IleValLeuCysTrpAlaGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe	226
652	TTAATGATTTGTCGTGCAGAGGGCTGTATTTTCAGAACAAATCTCTTATATCTTCTAC	711

Qy 571 GlnLysLeuLeuHisTrpValSerLeuLeu-----GlyGlnGlnProAsnAlaThrThr 588
 Db 1732 CGACAGTGTGTTTGTGGCTGGAATCTCTATTGGACACTTTACATCTCGAGCTCAAAAAA 1791
 Qy 589 ProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheVal 608
 Db 1792 ATA---AATACCATGAAGTCTTTTACTGCTGTTTGAGATGGAAGAGAGTCTTTGTA 1848
 Qy 609 ArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeu 628
 Db 1849 CAGTCAGCAATGAACCTAGGGAACAGATGACGTGTGGTTAAGGGTTATTCTGATTTT 1908
 Qy 629 IleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspVal 648
 Db 1909 ATTGTGTGCTGCTACTGCTTAAGCCATGGCTGTGCACGTACAGACTTCTCCATTCAGCT 1968
 Qy 649 LysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMet 668
 Db 1969 CAAATGTGCTGAATGAAGAG-----CTAGGCCAG 1998
 Qy 669 ArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHis 688
 Db 1999 AGGGGAAGACTACTCAT-CTTTGGCATCAATATGCTCTGTGTTCTTAAGGAAT 2052
 Qy 689 ProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThr 708
 Db 2053 AAAGATATCAAAACACTCAGATGAAGACACTATTTTCAATGACCACTGCTTTAAAAATC 2112
 Qy 709 LeuCysAlaLysLeuArgHisProThrCysValIleGlnThrLeuMetPheArgAsnAla 728
 Db 2113 TTTTATCTTATTGGAAGACTCCAGCTGATCTCTTAAACACTTGTGGCATATAATGTT 2172
 Qy 729 GlnIleThrProGlyValGlnHisLeuThrArgIleValMetAlaAsnArgLeuArg 748
 Db 2173 TCTTTTATGTGAT--AAGCGCTATTCTTGGAAATGATTCAAGCTACAAATTTGGAG 2229
 Qy 749 SerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAla 768
 Db 2230 GAATTTGTATCTCAGAGGCACATCTCTCCACAGCTGATGTGGAAATGTTGTGTGATATC 2289
 Qy 769 LeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHis 788
 Db 2290 TTGAACGAGGAGTGCACATAGCGTACTGGAATTAGTAATCTCTCCCTCTGTGAG 2349
 Qy 789 AlaCysThrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSer 808
 Db 2350 CATCTCTGGGACTACTTCTGATGTTCTTAGCGAGACAATCTCTGAGTACTTAAC 2409
 Qy 809 LeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgVal 828
 Db 2410 ATCAGCTACAATAACCTGAAGGTGAAGGCTCTCTGTAGAGCTCTGACTCTC 2469
 Qy 829 SerGlnCysAlaLeuGlnLysLeuLeuGluAspCysGlyIleThrAlaThrGlyCys 848
 Db 2470 CCAAACTCTGCTGCTGATCTAAGTTTGAAGCTCTCAACTGACTGTGCTGTGT 2529
 Qy 849 GlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsn 868
 Db 2530 AAGGACCTTGCTTCTTACTTACCGGATATAAATCCCTGCGGAGGATCAACCTGCGGAG 2589
 Qy 869 AsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCys 888
 Db 2590 AACTCTTGGGCTTCAGTGGATTTGTTGTGCTGTGTAAGGCTATGAAAGACCAACGCTG 2649
 Qy 889 SerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeu 908
 Db 2650 ACCTTGTATGAGTAAACTCGAATGCGTACTGTTGTATGCGATCCAGGAATTTCTA 2709
 Qy 909 Ala-Leu----- 910
 Db 2710 CTTTCTGAGATGAAGAAATAAATCTTGAGCATCGAAATGTTGTGTGAAGAGAGCA 2769
 Qy 911 -----AlaLeuMetGlyAsnSerTrpLeu---ThrHisLeuSerLeuSe 924

Db 2770 GAAATCAAGGCTGGACATCTCTGATGCAAGAACAAATGGATGATTGTGAACCTCAAACTACT 2829
 Qy 924 rMetAsnPro---ValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluPr 943
 Db 2830 TCTGAATGAAGAGAGAGGATCTTTATTGTTGCTCTATTGTTGTAATAATTACAGGTACT 2889
 Qy 943 oSerCysHisLeuGlnAspLeuGluLeu 952
 Db 2890 AACACTGCAGACAGCAGGATTCTTTTGTG 2917
 RESULT 10
 AK054264 3400 bp mRNA linear HTC 20-SEP-2003
 LOCUS Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
 DEFINITION full-length enriched library, clone:E330007A02 product:Weakly
 similar to PAN1, full insert sequence.
 AK054264 GI:26344106
 VERSION HTC; CAP trapper.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE Carninci,P. and Hayashizaki,Y.
 1
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2
 REFERENCE Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 1
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3
 REFERENCE Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yumoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuoka,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuyama,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 1
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 JOURNAL sequencing pipeline with 384 multicapillary sequencer
 MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
 PUBMED 11076861
 4
 REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 1
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 1
 TITLE Analysis of the mouse transcriptome based on functional annotation
 JOURNAL of 60,770 full-length cDNAs
 REFERENCE Nature 420, 563-573 (2002)
 6
 (bases 1 to 3400)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
 Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
 Okazaki,Y., Saito,R., Saito,H., Sakai,C., Sakai,K., Sakazume,N.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

[illegible]

```
Db 1165 CTCAAGGAAGTGAGATGCAATGCTGATCTATACCAAAATGGGCTCACTTCTCTACTGCAATGT 1224
Qy SerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerValAlaLeuProPhe 402
Db 1225 GGGATATTCTGCTCTGCTGGAGCTAGGATGAAGAGGGGGAAGACCTGCTCTGACA 1284
Qy 403 AenGlnThrLeuThrGlyLeuHisAlaPhe-----ValPheHisGlnLeuThr 419
Db 1285 TGTGACAGATACACCTCTATGTTCTGAACTTCTGTGCGAAGTCTTCTCATCAGAAACC 1344
Qy 420 ProArgGlyValValArgCys-----LeuAsnLeuGluGluArgValValLeu 436
Db 1345 -----TGTGAGGACCATCTCAATGAGGAATCCAAATCTTATTC 1383
Qy 437 LysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGly 456
Db 1384 AAAAAAATATGATCTGCTGCTGCCAATAGTCTCTTGGAAACAGGTGCCCATATCTGTGAA 1443
Qy 457 AspAspLeuMetValGlnGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsn 476
Db 1444 GAAGACTTCTTGACATTAAATTTGAACCTAAACAATTTGCATCCCATGGTGTGCAGGCAC 1503
Qy 477 IleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGln 496
Db 1504 ATCTCTTTTAAGGACAGTTCAGATACACACTGTCTCTTTTCATCTGCTCGGTATCCAG 1563
Qy 497 AspPheCysAlaAlaLeuTyrTyrValLeuGlu---GlyLeuGluIleGluProAlaLeu 515
Db 1564 CAACTCTGCTGGCTGCAATATATTGTTGAGGAGCTTGACAGGAAAG-CAAGGGTGTCTC 1622
Qy 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db 1623 TAAATATAGTATACAGAACATGCTTTCAAGGAGCAAGATTAATAAACCC----- 1673
Qy 536 IleHisSerLeuTrp-MetLysArgPheLeuPheGlyLeuValSerGluAspValArgAr 555
Db 1674 ---TGACCTGTCTGGGCTGTGCCATTTGTATTTGGCCCTCTGTAATGAGACTCGGTATTCA 1730
Qy 555 gProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHi 575
Db 1731 AGAGTTGAAGACTACTTTGGCTGCCAATATCAACTGAGGTCAAGAGAAATTCCTGGA 1790
Qy 575 eTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAl 595
Db 1791 GTGTGAGTCA-----GGGAAATAAACCCCTTACTATTATGATGATATGCAAGAAAT 1844
Qy 595 aPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPh 615
Db 1845 CTGTCTCTCTATGAATCTCAGAGAGGGGTTTGTGAAGGAAGCCATGGTCTCTT 1904
Qy 615 eGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLe 635
Db 1905 TGAGGACATCTCTTGGCATTTTGAACACGACACTGACCTCATTCATGCTTATTCCTGCT 1964
Qy 635 uGlnHisCysProTyrLeuArgLysIleArgValAspVal---LysGlyIlePheProAr 654
Db 1965 CAAGAACTCTCAGAACTTGCAGACAATGCTCCCTGGAAGTGAAGAGCGGCTTTCCAGA 2024
Qy 654 gAspGluSerAlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIle 674
Db 2025 GAATGTGTGCTGCACTAGAGTCCACAGTCAAGCATCAAAAGGTCCCGGATGAACAACGCAT 2084
Qy 674 eGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLe 694
Db 2085 GTTAATCTTCTGGACAGACTTTTGTGACACATTTAATTCAAATAAGAAACCTGGTATTCT 2144
Qy 694 uAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuAr 714
Db 2145 AGACATCCATGAAGAGCTCTCTCAACAGCTCTGCACTGGGAGATTCTTTGTGAAGAGTCC 2204
Qy 714 gHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyVa 734
```

RESULT 11

```
BM454123          1119 bp      mRNA      linear      EST 05-FEB-2002
LOCUS             AGENCOURT 6419582 NCI CGAP Ov44 Mus musculus cDNA clone
DEFINITION        IMAGE:5501420 5', mRNA sequence.
ACCESSION         BM454123
VERSION           BM454123.1 GI:18503163
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
Eukaryota; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1119)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
```

**JOURNAL
COMMENT**

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Aaron Hueh
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12145 row: 1 column: 21
 High quality sequence start: 3
 High quality sequence stop: 705.
 Location/Qualifiers
 1. .llil9
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5504420"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lbb="NCI_CGAP OV44"
 /notes="Organ: ovary, EMSG-treated; Vector: pCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dt. Average insert size 2.2 kb. Library
 constructed by Life Technologies. Note: this is a NCI_CGAP
 Library."

FEATURES

source

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5504420"
/lab_host="DHI0B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_OV44"
/notes="Organ: ovary, PMSG-treated; Vector: pCMV-SPORT6.1;
Site_1: BcorV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 2.2 kb. Library
constructed by Life Technologies. Note: this is a NCI_CGAP
Library."

ORIGIN

Alignment Scores:
Pred. No.: 9,95e-59 Length: 1119
Score: 687.00 Matches: 172
Percent Similarity: 64.55% Conservative: 52
Best Local Similarity: 49.57% Mismatches: 114
Query Match: 9.71% Indels: 13
DB: 12 Gaps: 6

```

ORIGIN

Alignment Scores:		
Pred. No.:	9.95e-59	Length:
Score:	687.00	Matches:
Percent Similarity:	64.55%	Conservative:
Best Local Similarity:	49.57%	Mismatches:
Query Match:	9.71%	Indels:
DB:	12	Gaps:
		6
US-10-066-521-6 (1-1344)	x	BM454123 (1-1119)

280	AsnAsnAspThrLysLeuCysLysAspTrrPalaGluLysGlnProPropheThrLeuile	299
Qy		
Db		
300	ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuileValThrValArgAsp	319
Qy		
Db		
320	ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly	339
Qy		
Db		
340	IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys	359
Qy		
Db		
360	ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro	379
Qy		
Db		
380	AlaValGlySerLeuileCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal	399
Qy		
Db		
400	AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr	419
Qy		
Db		
420	ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe	439
Qy		
Db		
440	CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu	459
Qy		
Db		
489	TGCATGATGGCAGCTGAAGAGAGTGGACCATGAGTCCGGTGTCTATGATGATGACCTG	548
Qy		
Db		

Qy	460	MetValGlnGlyLeuGlyGluSerGluLeuAArgAlaLeuPheHisMetAsnIleLeuLeu	479
Db	549	AAGAACTATAGCCCTAAAGGAGTCTGAGATCTTGGCCCTCTTTACATGAACATCCTTCTC	608
Qy	480	ProAspSerHis---CysGluGluTyThrPhePheHisLeuSerLeuGlnAspPhe	498
Db	609	CAGGTTGGCCACACACAGTGCAGAGTGTATGTTTTCTCCCACTTCAGCCTGCAGATTTC	668
Qy	499	CysAlaAlaLeuTyTrValLeuGluGlyLeu---GluIleGluProAlaLeuCysPro	517
Db	669	TTTGCTGCCCTTATATATGTTTTAGAAGGCTGAGGAATGGAATCAGCATTTTTGC---	725
Qy	518	LeuTyrValGluIlyThrIysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis	537
Db	726	--TTTCATTGAAACCAAGGAGCATCATGGAGGTGAAGAACTGAC---GACACTCGC	779
Qy	538	SerLeuTrpMetIysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu	557
Db	780	CTTCTCGGATGAAGCGCTTCTTATTGGCGCTCATGAACAAGGATATCTTTGAAGACTCTG	839
Qy	558	GluValLeuLeuGlyCysProValProLeuGlyValIysGln---LysLeuLeuHisTrpV	577
Db	840	GGAGGTCTCTTTGAATATCCCGTG---ATTCCACTGTTGACAAAGCTCCCACTGGG	896
Qy	577	alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH	597
Db	897	TCICTCTGATAGCTCAGCAGGC-CATGGCACCCGCCCAATGGACACCTG-GATGCCCTTCA	954
Qy	597	isCysLeuPheGluThrGlnAspIysGluPheValArgLeuAlaLeuAsnSerPheGlnG	617
Db	955	ATTGGCTATTGTATCT-CAGGATGAAAGTTGTGGCGGG-GCTCTCAAC---CTTTCCA	1009
Qy	617	luValTrpLeuProIle	622
Db	1010	AAGTGTGGCTGCTAAAT	1026

RESULT 12
CB228956
LOCUS
DEFINITION
CB228956
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

814 bp mRNA linear EST 10-FEB-2003
AGENCOURT 11501187 NICHD Rh. Ovi Macaca mulatta CDNA clone
IMAGE:6884760 5', mRNA sequence.
CB228956
CB228956.1 GI:28280534
EST.
Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 814)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: cgap@lml.llnl.gov
 Tissue Procurement: Dr. Elliot Spindel
 CDNA Library Preparation: CLONTECH
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L12CM3135 row: b column: 23
 High quality sequence row: 235.

FEATURES
SOURCE

```

source
i. ora
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884760"
/tissue_type="Ovary"

```

/lab_host="DH108 (phage-resistant)" /clone_lib="NICHD_Rh_Ov1" Note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I; Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.0-4.0 kb. Tissue pooled from pre-pubertal, post pubertal sn menopausal monkeys. Constructed by Clontech. Note: this is a NICHD Library."	
ORIGIN	
Alignment Scores:	
Pred. No.: 3,366-57	Length: 814
Score: 669.50	Matches: 140
Percent Similarity: 83.4%	Conservative: 11
Best Local Similarity: 77.35%	Mismatches: 27
Query Match: 9.46%	Indels: 3
DB: 14	Gaps: 1
US-10-066-521-6 (1-1344) x CB228956 (1-814)	
QY 718 CysLysIleGlnThrLeuMetPheArgAenAlaGlnIleThrProGlyValGlnHisLeu 737	
Db 3 TGCAGATACAGACCTCAGGCTTTAGGAACGCAGAGTTGCCTCTGCTGCACACCTC 62	
QY 738 TrpAqIleValMetAlaAsnArgAsnLeuArgSerLeuAenLeuGlyThrHisLeu 757	
Db 63 TGGAGAACCTCATTTGCCAACGTTAACTAGATCCCTCAATTTGGGAGGCCACCCCTG 122	
QY 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777	
Db 123 AGGAAGAGAGATGTAATGATGCGGTGTAAGCCTTAACACCCCAAAATGTTGTGGAG 182	
QY 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysIleLeuLysIleSerGlnIle 797	
Db 183 TCTTTGAGGCTGGATCGCTGTGATTGACCCAGCGCTGTACCTGAAGATCTCCCAATC 242	
QY 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817	
Db 243 CTTACGACCTCTCCCAACCTGAAGTCTCTGATCCTGGCAGGAAATGAGTGACAGACCAT 302	
QY 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837	
Db 303 GAGTGACCCCTCTCAGTGATGCTTTGAGGTCTCCCATGTCGCCCTCGAGAGCTGACA 362	
QY 838 LeuGlu-AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValse 857	
Db 363 CTGGAAGGACTGTGGCATCATGGCCACGGTTGCCAAGTCTGCGCCACACCTCGTCTG 422	
QY 857 rIenArgSerLeuThrHisLeuCysLeuSerAsnAsn-SerLeuGlyAsnGluGlyVala 877	
Db 423 CTACCGAACTAGACACATCTGTGCTATTCCACACCTATPCGAGGGGAAAGAGAGTGA 482	
QY 877 snLeuLeuCysArgSerMetArgLeuProHisCys--SerLeuGlnArgLeuMetLeu 895	
Db 483 ACCCGCTGTGATCATCAATTGAGGTTTCCCATGTAGTCTTCGCGAAAAATTTGATGCTG 541	
RESULT 13	
AK039531	
LOCUS	2730 bp mRNA linear HTC 19-SEP-2003
DEFINITION	Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched library, clone:A33005K17 product:similar to VASOPRESSIN RECEPTOR [Rattus norvegicus], full insert sequence.
ACCESSION	AK039531
VERSION	AK039531.1 GI:26087232
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253

```
/clone.lib="RIKEN full-length enriched mouse cDNA library"  
/dev stage="adult"  
1..2730  
/note="similar to VASOPRESSIN RECEPTOR [Rattus norvegicus]  
(SPTP|063035, evidence: FASTY, 82.2%ID, 100%length,  
match=1440)"
```

misc_feature

ORIGIN

Alignment Scores:

```
Pred. No.:      3,46e-56      Length:      2730  
Score:          669.50      Matches:      235  
Percent Similarity: 42.53%      Conservative: 132  
Best Local Similarity: 27.23%      Mismatches:  349  
Query Match:      9.46%      Indels:      149  
DB:              11          Gaps:       21
```

US-10-066-521-6 (1-1344) x AK039531 (1-2730)

```
QY 186 ArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 205  
DB 14 CAGCGCTGACCGTGGTCTCGAGGCGCCGCGAGCATCGGCAAAACCATGGCGCGCAAG 73  
QY 206 ArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhe 225  
DB 74 AAAATCCTCTATGACTGGCGACAGGTAAGCTATACACACGCCAGGTGGACTTTGCCCTC 133  
QY 226 PheLeuProValArgGlu---MetGlnArgLysLysGluSerSerValThrGluPheIle 244  
DB 134 TTCATGCCGTGCGGAGCTCTCGAGCGCGCGGAAAGCGCAGCTTGGCGGACCTGGTG 193  
QY 245 SerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArg 264  
DB 194 CTGACCATGCTCCCGACCGCGCTTGGCCAGTGAAGCGGATACTGGCGCAACCCCAATCGC 253  
QY 265 LeuLeuPheIleIleAspGlyPheAspLeuGlySerValLeuAsnAsnAspThrLys 284  
DB 254 CTACTGTTTCCTCGGAGCGTGCAGAGCGAGCTGCTACTTTACCATCCTCAGAGCCCACT 313  
QY 285 LeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArg 304  
DB 314 CCCTGCAAGACCCCTTTGGAGGCTACAAAGTGGCTTGAGAGTGTCTGAGCGGCTGCTGAGC 373  
QY 305 LysValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLys 324  
DB 374 CAGAGCTGTGTACAGAGCTCGCTTCTAGTGACTACACCCATCGCGCTCAGGGAGG 433  
QY 325 LeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGln 344  
DB 434 CTGACGGCGAGATTGTGCTGCCACAGTGCAGCAAAATACCGCGCTTCTCAGACAAAGAC 493  
QY 345 ArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArg 364  
DB 494 AAAAAAAGTATTTCTTCAAGTTCTTCGGGACGAGAGGAAGGACGAGCGCGCTTACCGC 553  
QY 365 AlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeu 384  
DB 554 TTCGTGAAGACAGACGAGACGCTGTTGCACTGTGCTTGTGGCCCTCTGCTGCTGGATC 613  
QY 385 IleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGln 404  
DB 614 GTGTGCACAGTGTGCAGCAGCAGCTAGAGCTGGCGCGGGGTCTCTCTCGTACCTCCAAAG 673  
QY 405 ThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValVal 424  
DB 674 ACCACTACTTCTGTGTACTCTCTTCAACACAGCATGTGTGAAGTCTGCGAGGCCACCAAT 733  
QY 425 ArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaVal 444  
DB 734 GGACCCCGGTTTCAGGAGAG-----CTGAGAACGCTGTGTCTGCTGGCCAGG 781  
QY 445 GluGlyValTrp---AsnArgLysSerValPheAspGlyAspLeuMetValGlnGly 463  
DB 782 GAGGGCATCCTGGATCATATAAAGCAAAATTTCTCAGAAAGAGGACCTGGAGAAATTCGAG 841
```

```
QY 464 LeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAsp---Ser 482  
DB 842 CTTTCGGGGTTCCCAAGTTTCAGACAATATTTCTCAACAAGAAAGAGATACACGAGGTGCTA 901  
QY 483 HisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeu 502  
DB 902 AAAACTGAGGTACCTACCGTTTCATTCACACAGCTTCACGAGGTCTTGGGTGCATTG 961  
QY 503 TyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLys 522  
DB 962 TCATACCTACTAGAACT---GAGCGAACTCCCGGAGCGCCCGCAGCGCGTGTGCAGAAG 1018  
QY 523 ThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTyrMetLys 542  
DB 1019 CTCCTGAATCT-----GACGCGGAGCTAGTGTGTCATCTTGCACTGACCCACC 1066  
QY 543 ArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGly 562  
DB 1067 CGATTCTCTTTGGCTCTGTAATACAGAGGGCTTCGTGACATTGGAAACCAATTTTGGC 1126  
QY 563 CysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGln 582  
DB 1127 TGTGTGGTCCAGCATCATGTGAAAGGACACCTCGCGTGGGTACAA-----GGACAG 1180  
QY 583 ---GlnProAsnAlaThrThrProGlyAspThr----- 592  
DB 1181 AGCACCCCAAGGGGCCACCCAGTAGGGGCAAAAAGAGACTGCTGAGCTGGAGGACATCGAG 1240  
QY 593 -----LeuAspAlaPheHis 597  
DB 1241 GAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1300  
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617  
DB 1301 TGCCTGTATGAGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1360  
QY 618 ValTrpLeuProIleAsnGln-----AsnLeuAspLeuIleAlaSerSerPheCysLeu 635  
DB 1361 ATAGTACTGGAGCGAGTCTGAGTTGACCCGCTTGAGCTTGAAGTCTGAAATTAATCTCGTG 1420  
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655  
DB 1421 CAGTGTGCCAGATGGCAGGCTCTGAGACTGCTGAGTGTGGGTGTGGCGGCGCAAG 1480  
QY 656 Glu----- 656  
DB 1481 GAGAAGAAAGAAAGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1540  
QY 657 -----SerAlaGluAlaCys 661  
DB 1541 AAACAAACCCAGCTCTCTTGTGCGTCCACTCTGTGAGACATGACTACCCCGAAATGT 1600  
QY 662 ProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAspPhe 681  
DB 1601 CATCTGAGTGTCTGATCTTGTCCACTGAGACTCCCTGATGCAGTTTGGCCGAGACCTT 1660  
QY 682 CysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeu 701  
DB 1661 TCCGAGGCGCTGAAGGTAGTCTCTGCGCTAAGGGAGTTGGGCTCTCTCCAGAGCCGCGCTC 1720  
QY 702 ThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGln 721  
DB 1721 ACCAACACAGCGCTGCGTTTACTGTGCGAAGGCTGGCTTGGCCCAAGTCCAGGTTGAAG 1780  
QY 722 ThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleVal 741  
DB 1781 ACCTCAGATGAGCTGCTGACCTCCAGAGAGGTGATCAATTACTTGTGTCATCGTGCTC 1840  
QY 742 MetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGluAsp 761  
DB 1841 CAGCAGAGCCAGTCTCCACCACTCTGGACCTCAGTGGCTGTGAGTGTGCTGCGTGCATTT 1900
```


Qy 762 ValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeu 781
Db 1901 GTGGAACTCTGTGTGCGACCCCTGAAGCACCCTAAATGCTCCCTAAAGACCTCAGTCTG 1960
Qy 782 AspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrSer 801
Db 1961 ACCTCTGGAGCTGAGTGAGACTCAGTGGGATCTTCAAGCTGTGAAGACATCAAG 2020
Qy 802 ProSerLeuLys-SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetPr 821
Db 2021 CCAGATCTGTCCATATATTCAAAATAAGCACACCTCAGCGCTCGGA----- 2072
Qy 821 oLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCy 841
Db 2073 -----AGGGTGAGCAGTGT----- 2087
Qy 841 sGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLe 861
Db 2088 -----CTTCTGAAGCTCCGGAGTCC 2107
Qy 861 uThrHis-----LeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLe 879
Db 2108 AGAACACGATGAAGATGCTGTAGCCCAAGTCCCTCTG-AAT-----ATGACCTCCCC 2160
Qy 879 uCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMet----- 894
Db 2161 ATTCAAAAGCAAGAGATGCCAGC-TTCTCCCTCCAGGCAAGTCTTTCAAGCCACAGAAG 2219
Qy 895 -----LeuAsnGlnCysHisLeuAspThrAlaGl 904
Db 2220 GCGAAGCAAAAGACCCAGTGTAGATACAGCGCTTGATAGCTCTCGATAGACCTGGCAGC 2279
Qy 904 Y-----CysGlyPheLeuAlaLeuAla---LeuMetGlyAsnSerTr 917
Db 2280 TGCCCTGCCTGGACACTCTGCGGTGTCTCTCCCTAGTCCCAACACCAAGTCCCT 2339
Qy 917 pleu-----ThrHisLeuSerLeuSerMetAsnProValGluAspAsnGly----- 932
Db 2340 CTTTCTGAGCCTACCATCTCTGCTCCCTGACGATATCCCAAGTAATAGAAGCATTCAG 2399
Qy 933 -----ValLysLeuLeuCysGluValMetArgGluProSerCy 945
Db 2400 CAGCTGACGGTGTCTGAAATGTTGTCACTATGTGTGATCTG-----GCCCCGACGTG 2453
Qy 945 sHisLeu 947
Db 2454 TCATATG 2460
RESULT 14
LOCUS AY407369
DEFINITION Mus musculus RNH gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY407369
VERSION AY407369.1 GI:39763340
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1371)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003).
JOURNAL
PUBMED 14671302
REFERENCE 2 (bases 1 to 1371)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
source 1..1371
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
gene <1..>1371
/gene="RNH"
/locus_tag="HCM2867"

ORIGIN

Alignment Scores:
Pred. NO.: 4.27e-54 Length: 1371
Score: 643.50 Matches: 150
Percent Similarity: 52.31% Conservative: 76
Best Local Similarity: 34.72% Mismatches: 197
Query Match: 9.10% Indels: 9
DB: 29 Gaps: 3

US-10-066-521-6 (1-1344) x AY407369 (1-1371)

Qy 659 GluAlaCysProValValProLeuTrp-----MetArgAspLysThrLeuIle 674
Db 49 GAGCTCTTCTCTGATCCAGCATACGAGTGGTTCAGGCTGGATGACTGTGGCTCACT 108
Qy 675 GluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeu 694
Db 109 GAAGTGGGTGCAAGACATCAGCTCAGCAGTCCAGCAACCTGCCCTGACAGAGCTC 168
Qy 695 AspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArg 714
Db 169 AGCCTTCGCACCAATGAAGTGGGTGATGGGTCTGTGGTCTCCAGGGCTCGAG 228
Qy 715 HisProThrCysValIleGlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGly 733
Db 229 ATCCCACTTGTAAATGCCAGAGCTGAGCCTTCAGAACTGCGGCTTGACGAGGCTGGC 288
Qy 734 ValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGly 753
Db 289 TGTGGGATCTGCTGCGCATGCTGCGCTCTTGTCTACCTGCGTGAGCTACATCTCAAT 348
Qy 754 GlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLys 773
Db 349 GATAACCTATGGGGAGCGCAGCGCTGAAGCTGCTCTGTGAAGGACTTCAGGACCCGAG 408
Qy 774 CysLeuLeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLys 793
Db 409 TCGCGTCTTGAGAAGCTTCAGTTGGATATCTGTAACCTCAGCTACAGCTCGGAGCCCC 468
Qy 794 IleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLys 813
Db 469 CTGCTTCAGTGTCTCAGGGTGAAGCTGACTTTAAGGAGCTAGTATTGAGCAACACGAC 528
Qy 814 ValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeu 833
Db 529 CTCCTCAACCTGGTGTCCGTATTCTGTCCAGGGGCTGAAGGACTTCGCTGTCAACTG 588
Qy 834 GlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSer 853
Db 589 GAGTCACCTCAAGCTGGAGAACTGTGGTATCAGCAGCACTCAAGCAAGGATCTGTGGCAT 648
Qy 854 AlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsn 873
Db 649 GTTGTGGCTCCCAAGCTTCGCTACAGAACTGGACTTGAGCAACACAGTAGGTAAAT 708
Qy 874 GluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeu 893
Db 709 GCAGGCATTCAGCACTGTGCCAGGAGTCTGTCTCCAGCTGTAAAGCTCAGGACTCTG 768

Qy 894 MetLeuAanGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMet 913
 Db 769 TGGCTCTGGATGTGATATCACTCAGAGGCGTGAAGACCTGTGCCGTGCTCAGA 828
 Qy 914 GlyAanSerTrpLeuThrHisLeuSerLeuSerMetAanProValGluAspAsnGlyVal 933
 Db 829 GCCAAGCAGAGCCTGAAGAACTCAGCCTTCTTCCATGAGCTGAAGGATGAGGTGCC 888
 Qy 934 LysLeuLeuGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuVal 953
 Db 889 CGGCTGCTGTGTGAGAGCTATTAGAGCTGGCTGTGCTGAGTGGAGTCACTGTGATAAAG 948
 Qy 954 LysCysHisLeuThrAlaAlaCysGluSerLeuSerCysValIleSerArgSerArg 973
 Db 949 ACCTGTAGCCTCAGAGTGCCTCTCTGCTCCCTTCTGCTCCGTGTGACCAAAAGCGT 1008
 Qy 974 HisLeuLysSerLeuAspLeuThrAspAanAlaLeuGlyAspGlyGlyValAlaAlaLeu 993
 Db 1009 TCTGTGTGGAGTGCATAGCAGCAACCATCTGGAGACGAGGAGTCCAGAGCTT 1068
 Qy 994 CysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGly 1013
 Db 1069 TCCAAGGCGCTGAGCCAGCCAGATACAGTGTGCTGAGCTTGTGTGGAGACTGTGAT 1128
 Qy 1014 LeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThr 1033
 Db 1129 GTGACAAATAGTGTGTGAGAGCCTTCCCAATGTCTGTGGCAACCGCAGCTGAGG 1188
 Qy 1034 SerLeuAanLeuValGlnAsnAspPheSerProLysGlyMetMetLysLeuCysSerAla 1053
 Db 1189 GAACCTGGACCTCAGTAACAACACTGGGGGCCGAGGTGTCTGCAACTGCTGAGAGC 1248
 Qy 1054 PheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeu-----TrpLysTrp 1069
 Db 1249 CTCAAACAGCCGAGCTGCACCTTTCAGCAGCTGCTGTGTATGACATTTACTGACAAAT 1308
 Qy 1070 GlnTyPrProValGlnIleArgLysLeuLeuGlu 1081
 Db 1309 GAGGTGGAAGAACACTTCGGGCGCTGGAGAGGAA 1344

RESULT 15
 BB555431
 LOCUS
 DEFINITION BB555431 RIKEN full-length enriched, 2 days pregnant adult female
 ovary Mus musculus cDNA clone E330019F16 3' similar to AF074018 Mus
 musculus ooplasmic specific protein (OP1) mRNA, mRNA sequence.
 BB555431
 BB555431.2 GI:16448559
 EST.
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 671)
 AUTHORS
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
 Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, O.,
 Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
 Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)
 TITLE
 JOURNAL
 COMMENT On Aug 1, 2000 this sequence version replaced gi:9641797.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp.

URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Kira, A.,
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
 and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
 Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES

Location/Qualifiers
 1..671
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="E330019F16"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="2 days pregnant adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 2 days pregnant
 adult female ovary"
 /notes="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGCGCGCACTGAGTCTTTTCTTTTCTTTT 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGAGATTCCTGAGTAAATAATCCCTCCCTCCCTCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda FLC 1."

ORIGIN

Alignment Scores:
 Pred. No.: 6,33e-54 Length: 671
 Score: 635.00 Matches: 121
 Percent Similarity: 80.87% Conservatives: 27
 Best Local Similarity: 86.12% Mismatches: 35
 Query Match: 8.99% Indels: 1
 DB: 10 Gaps: 0

US-10-066-521-6 (1-1344) x BB555431 (1-671)

Qy 922 SerLeuSerMetAanProValGluAspAsnGlyValLysLeuLeuCysGluValMetArg 941
 Db 6 AGCCTGACCATGAACCCCGTAGGGATGTGTGCAATGAAGCTACTGTGTGAAGCTTTAAAG 65

```
QY 942 GluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCys 961
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 GAACCTACTTGTAC-CTTCAGAACTGGAACTAGTGGACTGCCAACTCACACAGAACTGC 124
QY 962 CysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThr 981
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 TCGGAGGACCTGGCTGTATGATCACAACAAGCACTTAAAGTTGGATCTTGGT 184
QY 982 AspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsn 1001
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 AACAAAGCCCTGGGTGACAAAGAGTCTATACCCCTGTGTGAGGGACTGAGCAAGTAGC 244
QY 1002 SerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAla 1021
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 AGCTCCCTGAGGAGACTTGGGTGGGGGCATGTAAGTTGACTTCCAATTGCTGTGAGGCA 304
QY 1022 LeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsn 1041
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 TTGTCTATTGGCCATCTCTTGCAACCCCTCACCTGAACAGCCCTAAACCTGGTGAAGAATGAC 364
QY 1042 PheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeu 1061
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 TTCAGTACATCGGGATGTTGAAGCTGTGCTCTGGCTTCCAATGCCCTGTCTCTAACCTG 424
QY 1062 GlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuGluGlu 1081
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 GGGATAATTGGCCTGTGGAAGCAGGAGTACTATGCCCGAGTGAGAAAGACAGCTGGAGGA 484
QY 1082 ValGlnLeuLysProArgValIleAspGlySerTrpHisSerPheAspGluAsp 1101
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
485 GTTGAGTTTGCAAGCCCCACCGTGGTGATGATGGTGATGGTATGCTAGTGATGAAGAT 544
QY 1102 AspArgHis 1104
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
545 GACCGAAAC 553
```

Search completed: July 16, 2004, 06:23:48
Job time : 7508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 09:47:35 ; Search time 198 Seconds
(without alignments)
3766.939 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAGLEGVLVS 1344

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spoel/US1006521/runat 13072004 122318 11560/app_query.fasta_1.1543
-DB=Issued_Patents_NA -QFMT=fascap -SUFFIX=Oligo.rni -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1006521@cgn 1 1.107 @runat 13072004 122318 11560 -NCPUS=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	0.6	29	4	US-09-100-703A-19
2	8	0.6	423	4	US-09-621-976-16796
3	8	0.6	426	4	US-09-621-976-16797
C 4	8	0.6	549	4	US-09-252-991A-5717
5	8	0.6	618	4	US-09-621-976-788
C 6	8	0.6	805	3	US-08-961-083-207
C 7	8	0.6	805	4	US-09-536-784-207
8	8	0.6	855	4	US-09-489-039A-2718
9	8	0.6	891	4	US-09-252-991A-5722
C 10	8	0.6	944	4	US-09-262-856A-13
C 11	8	0.6	958	2	US-08-632-514C-9
C 12	8	0.6	958	3	US-09-188-177-9

C 13	8	0.6	1117	4	US-08-858-207A-40	Sequence 40, Appl
C 14	8	0.6	1126	4	US-08-961-527-13	Sequence 13, Appl
C 15	8	0.6	1176	4	US-09-491-577-75	Sequence 75, Appl
C 16	8	0.6	1272	4	US-09-489-039A-6621	Sequence 6621, Ap
17	8	0.6	1278	4	US-09-252-991A-5750	Sequence 5750, Ap
18	8	0.6	1299	4	US-09-489-039A-6764	Sequence 6764, Ap
19	8	0.6	1566	4	US-09-833-381-1975	Sequence 1975, Ap
20	8	0.6	1599	4	US-09-252-991A-13921	Sequence 13921, A
C 21	8	0.6	1622	3	US-09-180-939-2	Sequence 2, Appli
C 22	8	0.6	1750	4	US-09-262-856A-7	Sequence 7, Appli
C 23	8	0.6	2155	4	US-09-371-338-8	Sequence 8, Appli
C 24	8	0.6	2280	4	US-09-252-991A-13535	Sequence 13535, A
25	8	0.6	2866	4	US-09-724-864-30	Sequence 30, Appl
C 26	8	0.6	3318	4	US-09-540-236-1379	Sequence 1379, Ap
27	8	0.6	3336	4	US-09-717-364A-26	Sequence 26, Appli
C 28	8	0.6	3393	3	US-09-104-324B-1	Sequence 1, Appli
C 29	8	0.6	3393	3	US-09-162-713-1	Sequence 1, Appli
C 30	8	0.6	3446	4	US-09-332-714-13	Sequence 13, Appl
C 31	8	0.6	3665	4	US-09-976-594-722	Sequence 722, App
32	8	0.6	4086	4	US-09-702-705-1801	Sequence 1801, Ap
33	8	0.6	4086	4	US-09-736-457-1801	Sequence 1801, Ap
34	8	0.6	4086	4	US-09-671-325-1801	Sequence 1801, Ap
C 35	8	0.6	5228	4	US-09-428-711A-15	Sequence 15, Appl
C 36	8	0.6	5418	3	US-09-180-939-1	Sequence 1, Appli
37	8	0.6	5578	1	US-08-081-610-2	Sequence 2, Appli
C 38	8	0.6	5759	4	US-09-898-361-3	Sequence 3, Appli
39	8	0.6	8561	3	US-09-112-450-3	Sequence 3, Appli
40	8	0.6	8561	4	US-09-419-291A-3	Sequence 3, Appli
41	8	0.6	8876	4	US-08-961-527-170	Sequence 170, App
42	8	0.6	26016	4	US-09-326-480A-1	Sequence 1, Appli
43	8	0.6	26270	4	US-09-717-364A-1	Sequence 1, Appli
44	8	0.6	40085	4	US-08-311-731A-26	Sequence 26, Appl
45	8	0.6	50000	4	US-09-146-053-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-100-703A-19
; Sequence 19, Application US/09100703A
; Patent No. 6610300
; GENERAL INFORMATION:
; APPLICANT: SEGERS, Ruud PAM
; APPLICANT: WATERFIELD, Nicolas R
; APPLICANT: FRANDSEN, Peer L
; APPLICANT: WELLS, Jeremy M.
; TITLE OF INVENTION: COLISTRIDIMUM PERFRINGENS VACCINE
; FILE REFERENCE: 97288 US
; CURRENT APPLICATION NUMBER: US/09/100,703A
; CURRENT FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: EP97201888.1
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 19
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Clostridium perfringens
US-09-100-703A-19

Alignment Scores:
Pred. No.: 8.25 Length: 29
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-100-703A-19 (1-29)

Oy 1023 SerleuAlaLeuSerCysAenArg 1030

Db 6 TCTCTAGCTTTGAGCTGTAATAGA 29

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5717
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5717

Alignment Scores:
Pred. No.: 157      Length: 549
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-252-991A-5717 (1-549)
QY 849 GlnSerLeuAlaSerAlaLeuVal 856
Db 109 CAGAGCCTGGCATCGGGCTGGTG 86

RESULT 5
US-09-621-976-788
; Sequence 788, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 788
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 306..617
US-09-621-976-788

Alignment Scores:
Pred. No.: 176      Length: 618
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-621-976-788 (1-618)
QY 904 GlyCysGlyPheLeuAlaLeuAla 911
Db 26 GCCTGTGGGTTTGGCTCTTGCA 49

RESULT 6
US-08-961-083-207/c
; Sequence 207, Application US/08961083
; Patent No. 6159469

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16796
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16796

Alignment Scores:
Pred. No.: 122      Length: 426
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-621-976-16797 (1-426)
QY 904 GlyCysGlyPheLeuAlaLeuAla 911
Db 55 GCCTGTGGGTTTGGCTCTTGCA 78

RESULT 4
US-09-252-991A-5717/c
; Sequence 5717, Application US/09252991A
; Patent No. 6551795
```



```
;
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-207

Alignment Scores:
Pred. No.: 230 Length: 805
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 3 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-961-083-207 (1-805)
Qy 918 LeuThrHisLeuSerLeuSerMet 925
Db 801 TTAACCCACTTATCATTCATG 778

RESULT 7
US-09-536-784-207/c
; Sequence 207, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 207:
; US-09-536-784-207

Alignment Scores:
Pred. No.: 230 Length: 805
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-536-784-207 (1-805)
Qy 918 LeuThrHisLeuSerLeuSerMet 925
Db 801 TTAACCCACTTATCATTCATG 778

RESULT 8
US-09-489-039A-2718
; Sequence 2718, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2718
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-2718

Alignment Scores:
Pred. No.: 244 Length: 855
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-489-039A-2718 (1-855)
Qy 830 GlnCysAlaLeuGlnLysLeuIle 837
Db 689 CAGTGGCGCTTACAAAACTATC 712

RESULT 9
US-09-252-991A-5722
; Sequence 5722, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; CORRESPONDENCE ADDRESS:
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5722
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5722

Alignment Scores:
Pred. No.: 254 Length: 891
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-252-991A-5722 (1-891)

Qy 849 GlnSerLeuAlaSerAlaLeuVal 856
Db 316 CAGAGCCTGGCATGGCGCTGGTG 339

RESULT 10
US-09-262-856A-13/c
; Sequence 13, Application US/09262856A
; Patent No. 6333164
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, Kazutosh
; APPLICANT: MIZUTANI, Shigetoshi
; APPLICANT: ENDO, Masahiro
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 1422-372P
; CURRENT APPLICATION NUMBER: US/09/262,856A
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-262-856A-13

Alignment Scores:
Pred. No.: 269 Length: 944
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-262-856A-13 (1-944)

Qy 797 IleLeuThrThrSerProSerLeu 804
Db 713 ATCTTGACCACTTCACCATCTT 690

RESULT 11
US-08-632-514C-9/c
; Sequence 9, Application US/08632514C
; Patent No. 5834234
; GENERAL INFORMATION:
; APPLICANT: GALLO, Gregory J.
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,177
; FILING DATE: 2
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
```

```
; APPLICANT: GALLO, Gregory J.
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/632,514C
; FILING DATE: 29-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.188
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8459
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-632-514C-9

Alignment Scores:
Pred. No.: 273 Length: 958
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 2 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-632-514C-9 (1-958)

Qy 1218 LeuGlyArgGluLeuSerSerArg 1225
Db 583 CTTGGCGAGAGCTGTCATCCCGG 560

RESULT 12
US-09-188-177-9/c
; Sequence 9, Application US/09188177
; Patent No. 6057132
; GENERAL INFORMATION:
; APPLICANT: GALLO, Gregory J.
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,177
; FILING DATE: 2
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
```

```

; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.188
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8459
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-188-177-9
Alignment Scores:
Pred. No.: 273 Length: 958
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 3 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-188-177-9 (1-958)
Qy 1218 LeuGlyArgGluSerSerArg 1225
Db 583 CTTGGCGAGAGCTGTCATCCGG 560

RESULT 13
US-08-858-207A-40/C
; Sequence 40, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1117 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

US-08-858-207A-40
Alignment Scores:
Pred. No.: 319 Length: 1117
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-858-207A-40 (1-1117)
Qy 918 LeuThrHisLeuSerLeuSerMet 925
Db 192 TTAACCCACTATCATTCATCATG 169

RESULT 14
US-08-961-527-13/c
; Sequence 13, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-13
Alignment Scores:
Pred. No.: 321 Length: 1126
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-961-527-13 (1-1126)
Qy 775 LeuLeuGluSerLeuArgLeuAsp 782
Db 1088 CTGCTGGAGTCACTTCGACTGGAT 1065

RESULT 15

```

```

US-09-491-577-75/c
; Sequence 75, Application US/09491577
; Patent No. 6610511
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: Carlson, John R.
; APPLICANT: Kim, Hunhyong
; APPLICANT: Clyne, Peter J.
; APPLICANT: Warr, Coral G.
; TITLE OF INVENTION: No. 6610511el Family of Odorant Receptor Genes in Drosophila
; FILE REFERENCE: 44574-5061-US
; CURRENT APPLICATION NUMBER: US/09/491,577
; CURRENT FILING DATE: 2000-01-25
; EARLIER APPLICATION NUMBER: US 60/117,132
; EARLIER FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1176)
; OTHER INFORMATION: DORLU 9.1
US-09-491-577-75

```

```

Alignment Scores:
Pred. No.: 336 Length: 1176
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

```

US-10-066-521-6 (1-1344) x US-09-491-577-75 (1-1176)

```

Qy 34 LysLysLysSerSerGluSerThr 41
Db 644 AAAAAGAGAGCGAGCGAGTCCACG 621

```

Search completed: July 16, 2004, 18:08:22
Job time : 209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 13:36:00 ; Search time 1299 Seconds
(without alignments)
5047.629 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3190992 seqs, 2439311697 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6373542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/usPTO_spool/US10066521/runat_13072004_122319_11640/app.query.fasta_1.1543
-DB=Published_Applications_NA -QFM=fastap -SUFFIX=Oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10066521 @CIGN 1 1 876 @runat_13072004_122319_11640
-NCPU=6 -ICPU=3 -NO_WRAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:**
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:**
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:**
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:**
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:**
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:**
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:**
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:**
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:**
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:**
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:**
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:**
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:**
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:**
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:**
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:**
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:**
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:**
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
------------	-------	--------------	-------	-------------

RESULT 1			
US-10-124-498-5	1344	100.0	US-10-124-498-5
; Sequence 5, Application US/10124498			
; Publication No. US20030017983A1			
; GENERAL INFORMATION:			
; APPLICANT: Bertin, John			
; APPLICANT: Wang, Weiye			
; APPLICANT: Blatcher, Maria			
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR			
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF			
; FILE REFERENCE: 07334-367001			
; CURRENT APPLICATION NUMBER: US/10/124,498			
; CURRENT FILING DATE: 2002-04-17			
; PRIOR APPLICATION NUMBER: 10/066,521			
; PRIOR FILING DATE: 2002-01-31			
; PRIOR APPLICATION NUMBER: 60/318,645			
; PRIOR FILING DATE: 2001-09-10			
; PRIOR APPLICATION NUMBER: 60/265,231			
; PRIOR FILING DATE: 2001-01-31			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 4035			

ALIGNMENTS

1	1344	100.0	4035	13	US-10-124-498-5	Sequence 5, Appli
2	1344	100.0	4035	15	US-10-066-521-5	Sequence 5, Appli
3	1006	74.9	3489	13	US-10-416-642-3	Sequence 3, Appli
4	1006	74.9	3926	15	US-10-216-645-1	Sequence 1, Appli
5	988	73.5	3830	15	US-10-216-645-3	Sequence 3, Appli
6	968	72.0	3226	13	US-10-092-900A-347	Sequence 347, App
7	616	45.8	3900	12	US-10-677-943-23	Sequence 23, Appl
8	616	45.8	3900	13	US-10-399-443-23	Sequence 23, Appl
9	385	28.6	1157	12	US-10-677-943-1	Sequence 1, Appli
10	385	28.6	1157	13	US-10-399-443-1	Sequence 1, Appli
11	196	14.6	1075	12	US-10-677-943-3	Sequence 3, Appli
12	196	14.6	1075	13	US-10-399-443-3	Sequence 3, Appli
13	57	4.2	2099	13	US-10-027-632-258159	Sequence 258159,
14	57	4.2	2099	16	US-10-027-632-258159	Sequence 258159,
15	12	0.9	509	15	US-10-029-386-10957	Sequence 10957, A
16	12	0.9	1704	15	US-10-029-386-24660	Sequence 24660, A
17	12	0.9	3108	10	US-09-965-621-23	Sequence 23, Appl
18	12	0.9	3108	16	US-10-407-866-23	Sequence 23, Appl
19	12	0.9	3186	13	US-10-124-498-17	Sequence 17, Appl
20	12	0.9	3186	15	US-10-066-521-17	Sequence 17, Appl
21	12	0.9	3218	16	US-10-407-866-67	Sequence 67, Appl
22	12	0.9	3447	12	US-10-677-943-5	Sequence 5, Appli
23	12	0.9	3447	13	US-10-399-443-5	Sequence 5, Appli
24	12	0.9	3466	16	US-10-108-260A-718	Sequence 718, App
25	12	0.9	4931	15	US-10-028-374-1	Sequence 1, Appli
26	12	0.9	4931	15	US-10-183-770-1	Sequence 1, Appli
27	10	0.7	610	13	US-10-424-599-139217	Sequence 139217,
28	9	0.7	240	15	US-10-029-386-16506	Sequence 16506, A
29	9	0.7	270	12	US-09-922-293-1943	Sequence 1943, Ap
30	9	0.7	339	17	US-10-437-963-46914	Sequence 46914, A
31	9	0.7	440	13	US-10-424-599-35497	Sequence 35497, A
32	9	0.7	564	15	US-10-029-386-2806	Sequence 2806, Ap
33	9	0.7	578	10	US-09-764-891-1745	Sequence 1745, Ap
34	9	0.7	720	13	US-10-027-632-150437	Sequence 150437,
35	9	0.7	720	16	US-10-027-632-150437	Sequence 150437,
36	9	0.7	1026	17	US-10-437-963-16592	Sequence 16592, A
37	9	0.7	1033	13	US-10-412-699B-1129	Sequence 1129, Ap
38	9	0.7	1033	16	US-10-374-780A-676	Sequence 676, App
39	9	0.7	1296	13	US-10-412-699B-1128	Sequence 1128, Ap
40	9	0.7	1296	16	US-10-374-780A-675	Sequence 675, App
41	9	0.7	1842	13	US-10-424-599-80385	Sequence 80385, A
42	9	0.7	1907	13	US-10-027-632-100254	Sequence 100254,
43	9	0.7	1907	16	US-10-027-632-100254	Sequence 100254,
44	9	0.7	2596	13	US-10-653-595-99	Sequence 99, Appl
45	9	0.7	2596	13	US-09-397-945-99	Sequence 99, Appl

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4032)
; US-10-124-498-5

Alignment Scores:
Pred. No.: 0 Length: 4035
Score: 1344.00 Matches: 1344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-124-498-5 (1-4035)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAGGAAGAAATTTACAGACATTTCAAGGAATTTACTAAAGAAGAAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluLeuLeuAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCCACAGTTTGAATTCAGAAATGCAACGTTGGAATGTCGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIlePheGlu 80
Db 181 CTCCTTCATGATGATTATGGAGCATCGCTGGCTGGGCTAGCTCCATTAGGATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGln 100
Db 241 AACATGAACCTCGCAACCTCTCGAGAAAGGACCGGATGACATGAAAGAAATTTCAAA 300
Qy 101 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlnIleSerGlnAla 120
Db 301 GCTATGGAAACAAAGAGTGGCCACAGCAGACAGACAGAAAGAAATTTCAAGCT 360
Qy 121 MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlyHisGlyAspThr 140
Db 361 ATGGAACAAAGAGTGGCCACAGCAGCAGACAGACAGAAAGAAATTTCAAGCT 420
Qy 141 TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArgSer 160
Db 421 TGGGACTACAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCTAGT 480
Qy 161 PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer 180
Db 481 TTTGAAAACACTGCTGCTGACTGGCCGGAATGAAACGTTGGCTGTTGATTTCA 540
Qy 181 AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLys 200
Db 541 GACCGTGGGGCTTCGGGCTTCGACCGTGTCTGACAGGAAAGTCAAGAAATTTGGGAA 600
Qy 201 SerAlaLeuAlaArgArgIleValLeuLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMet 220
Db 601 TCGGCTCTAGCCAGGAAGATCGTGTGTCTGGCGCAAGGTGGCTCTACAGGGAATG 660
Qy 221 PheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSerVal 240
Db 661 TTCCTCTACGCTTCTCTCCCTCCCGTTAGAGAGATGACGCGGAAGAGGAGAGAGTGC 720
Qy 241 ThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSer 260
Db 721 ACAGATTCATCTCCAGGGAGTGGCCAGACTCCACAGGCTCCGGTGGACGGAGATCATGTCC 780
Qy 261 ArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsn 280
Db 781 CGACCAGAAAGGCTGTGTTTCATCATTCACGCTTTCGATGACCTGGGCTCTGCTCTCAAC 840
```

```
Qy 281 AsnAspThrLysLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuLeuArg 300
Db 841 AATGACACAAAGCTCTGCAAGAGCTGGCTGGAAGCAGCGCTCGTTTCCACCTCATACGC 900
Qy 301 SerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspVal 320
Db 901 AGTCTGCTGAGGAAGGTCTCTGCTCCCTGAGTCTCTCTGATCGTCAACCGTCAGACGCTG 960
Qy 321 GlyThrGluLysLysSerSerGluValValSerProArgTyrLeuLeuValArgGlyIle 340
Db 961 GGCACAGAGAAGCTCAAGTCAGAGGTCTGTCTCCCGTTTACCTGTTAGTAGGAATC 1020
Qy 341 SerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLysThr 360
Db 1021 TCGGGGAACAAAGAAATCCACTTCTTGACGGCGGATTTGTTGAGCATCAGAAGACA 1080
Qy 361 GlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla 380
Db 1081 CAAGGGTGTGTCGATCATGAACACCGTGAAGTGTCTCGACCAAGTGCAGAGTGCCTGCC 1140
Qy 381 ValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla 400
Db 1141 GTGGGCTCTCTCATCTGCTGGGCTGCTGAGCTGAGGACGCTGGTGGGGAGAGCGTGC 1200
Qy 401 ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro 420
Db 1201 CCCTTCAACCAACGCTCACAGGCTGACGGCTGTCGCGCTTTTGTGTTTCATCAGCTCAC 1260
Qy 421 ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys 440
Db 1261 CGAGGCTGTGTCGGGCTGCTCTCAATCTGGAGGAAGAGTGTCTCTGTAAGCGCTCTGC 1320
Qy 441 ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet 460
Db 1321 CGTATGCTCTGGAGGAGTGTGGAATAGGAAGTCAAGTGTGTTGATGTGACGACCTCAT 1380
Qy 461 ValGlnGlyLeuGlyLysSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuPro 480
Db 1381 GTTCAAGGACTCGGGAGTCTGAGCTCCGTCCTGTTTCATGAAATCTCTCTCTCCCA 1440
Qy 481 AspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla 500
Db 1441 GACAGCCACTGTGAGGAGTACTACACCTCTCCACCTCAGTCTCCAGGACTTCTGTGTC 1500
Qy 501 AlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrVal 520
Db 1501 GCCTGTGACTACGTGTGAGGGGCTGGAATCGAGCCAGCTCTCTGCCCTCTGTACGTT 1560
Qy 521 GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrp 540
Db 1561 GAGAAGCAAAAGAGTCCATGGAGCTTTAAACAGCAGGCTTCCATATCCACTCGCTTTGG 1620
Qy 541 MetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeu 560
Db 1621 ATGAAGCGTTTCTGTTGGCTCGTGAGCGAAGCGTAAGAGGCGCCACTGGAGGTCCTG 1680
Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCCGTCTCCCTGGGGGTGAAGCAGAAGCTTCTGCACCTGGGTCTCTCTGTTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTGAGCAGGCTTAATGCCACCCAGGAGACACCTTGGACGCTTCCACTGTCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTTGTGCTTGGCATTAAACAGCTTCCAAAGAGTGTGCTT 1860
Qy 621 ProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CCGATTAAACAGAACCTGGACTTGTAGAGCATCTTCTCTCTGCCCTCCAGCACTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAla 660
```


Db 1921 TTCCGGAATAATTCGGGTGGATGTCAAAGGGATCTTTCCAAAGAGATGAGTCCGCTGAGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAsp 680
Db 1981 TGTCTCTGTGGTCCCTCTATGGATGGGATAGAGACCTTCATTGAGGACAGTGGGAAGT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 2041 TTCTGCTCCATGCTTGGCACCACCACACACCTTGGCGGAGCTGGACCTGCGCAGCAGCATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysAlaVal 720
Db 2101 CTGACAGACGGCGGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCACTCCCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCTCTGATGTTTAGAATAGCACAGATTACCTCTGGTGTGCAGCACCTCTGGAGAAATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCACCACCTGAAGGAAGAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArg 780
Db 2281 GATGTAAAGGATGGCGTGTGAAGCCTTAAACACACCCAAATGTTTGTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGGATTGCTGGGATTGACCCATGCCCTGTACCTGGAAGATCTCCCAATCCTTACGACC 2400
Qy 801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCAGGGATATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTCAGTGTAGTCTCCAGAGTCTCCAGTGCAGCTGCAGAGAGCTGATATCTGAGGAC 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGCCAGGCTGGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLys 880
Db 2581 TTGACACACCTGTGCTATCCAAACACAGCCTGGGGAGCAAGGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGGCCACCCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACAGGCTGGCTGTGGTGTCTTCTGCATCTGGCTTATGGGTAACTCAAGCTGACGCAC 2760
Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLysCysGluValMet 940
Db 2761 CTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGGAGGTCAATG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAla 960
Db 2821 AGAGAACAATCTTGTATCTCCAGGACCTGGAGTTGGTAAAGTGTATCTCACCGCCCGC 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGTCTGTGAGGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCTGATATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAAATGCCCTGGGTGACGGTGGGTGTGCTGGCTGTGGAGGAGCTGAGCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGlu 1020

RESULT 2

US-10-066-521-5

; Sequence 5, Application US/10066521

Db 3001 AACAGTGTCTTGACAGAGACTCGGGTTGAAGGCATGTGGAGTACTTCTGATTGCTGTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCATCTCTCTTGGCCCTTTCCTGCACCGGCATCTGACCAGTCTAAACCTGTGTCAGNAT 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCTTTGGCTGTGCCACGCTCTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeuGlu 1080
Db 3181 TTACAGATAAATTTGGGCTGTGGAAATGGCAGTACCCTGTGCAATAAGGAAGCTCTGGAG 3240
Qy 1081 GluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAAGTGAGACTACTCAAGCCCGAGTCTGTAATTCACCGGTAGTTGGCATTCTTTTGTGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTCCGGCTCCCTGAAAGCCGGCATGGCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysIlySargValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTGTCTGGGGGATGAACCCAGACAGAGAAGAGCGTGTGTCGCTTCGCTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyGluSer 1160
Db 3421 TTCAGAGCAGTACACCATTTGCCAGTCTCTCTGCTGGCCACGGCAATTTGGTAGTCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGAGTTGCACACAGTGGAGCAGAGCTCCCGCAACCCATGGCAGCAGCAACACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTrpSerGlyAlaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTGAGTGTGGATATTCCGAGCCTGGTCTGAAACTGTCTGAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisGlyGlyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGGCTTGATCCCAACAGTCTGATCATGACACGAGAGTATGGCTGTGCTACTAGGAGGA 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGCTTGTGTCCAAACAGTGTCTGATGACCACAGCGGTGTCTCTGGT 3720
Qy 1241 HisTrpGluArgLeuGlySerArgGlyTrpCysLeuAsnSerAlaAspAspHisSerGly 1260
Db 3721 CACTGGAGCGGCTGGGCTCTTAGGGGCTGGTGTCTTAACAGTGTCTGATGACCACAGCGGT 3780
Qy 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAspAsp 1280
Db 3781 GTGTCTGTGTCATCTGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCAACAGTGTCTGATGAC 3840
Qy 1281 HisSerGlyValAlaTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSer 1300
Db 3841 CACAGCGTGTGGCTCTGCTCACTGGAGCGGCTGGGCTCGAGGGGCTTGTGTCTCAACAGT 3900
Qy 1301 AlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuVal 1320
Db 3901 GCTGATACACACGCGGTGTCTCTGGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTG 3960
Qy 1321 SerAsnSerAlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlu 1340
Db 3961 TCCAAACAGTGTGATGACCACAGCGGTGTCTCTGGTCACTTGGAGCGGCTGGGCTCGAG 4020
Qy 1341 GlyLeuValSer 1344
Db 4021 GGGCTGGTGTCT 4032

Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-334001
; CURRENT APPLICATION NUMBER: US/10/066,521
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4032)
US-10-066-521-5

Alignment Scores:

Pred. No.:	0	Length:	4035
Score:	1344.00	Matches:	1344
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-10-066-521-6 (1-1344) x US-10-066-521-5 (1-4035)

Qy	1	MetGluGlyAspLysSerLeuThrPheSerTyrGlyLeuGlnTrpCysLeuTyrGlu	20
Db	1	ATGGAGAGAGCAATCGCTCACCTTCCAGCTACGGCTGCAATGGTGTCTATGAG	60
Qy	21	LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysSerSerGluSer	40
Db	61	CTAGACAAGGAAGAAATTTACAGACATTCAGGAATTTACTAAAGAAAGAAATTTCA	120
Qy	41	ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAenValGluCysLeuAlaLeu	60
Db	121	ACCAATGCTCTATTCCACAGTTTGAATTCAGAAATCGCAACGTGGAATGTCGGC	180
Qy	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrSerIleSerIlePheGlu	80
Db	181	CTCTGCAATGAGTATTATGGAGCATCGCTGGCTGGGCTAGCTCCATAGCAATCTTGA	240
Qy	81	AenMetAenLeuAthrGluLeuSerGluLysAlaArgAspMetLysLysIleSerGln	100
Db	241	AACATGAACCTCGCAACCTCTCGAGAAGGACCGGATGACATGAAAAAATTTCA	300
Qy	101	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGlnAla	120
Db	301	GCTATGGAACCAAGAGGTGCCACAGCAGCAGACAGAAACAAAGAAATTTCA	360
Qy	121	MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlyHisGlyAspThr	140
Db	361	ATGGACCAAGAGGTGCCACAGCAGCAGACAGACAGACCAAGACATGGAGGTGAC	420
Qy	141	TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgSer	160
Db	421	TGGACTACAAAGAGTCAGCTGATGACCAAAATTCGCTGAGGAGGAGGATGAC	480
Qy	161	PheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer	180
Db	481	TTTGAACCACTGCTGCTGATGCTGGCCGGAATGCAACGTTGGCTGGTCTTTG	540
Qy	181	AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLys	200

Db	541	GACCGTGGGGCTTCGGGCTTCGACCGTGGTTCGACCGAAAGTCAGCAATGGGAAA	600
Qy	201	SerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMet	220
Db	601	TCGGCTCTAGCCAGAGGATCGTGTGCTGGCGCAAGGTGACTCTTACAGGGAATG	660
Qy	221	PheSerTyrValPhePheLeuProValArgGluMetClnArgLysLysGluSerSerVal	240
Db	661	TTCTCTACGTCTTCTTCTCCCGGTAGAGATGCGCGAAGAGAGAGACGAGTGC	720
Qy	241	ThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSer	260
Db	721	ACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACGGAGATCATGTCC	780
Qy	261	ArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsn	280
Db	781	CGACCAAGAAAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGCT	840
Qy	281	AsnAspThrLysLysCysLysAspTTPAlaGluLysGlnProProPheThrLeuIleArg	300
Db	841	AATGACACAAAGCTCTGCAAGAGCTGGGCTGGAAGCAGCTCCGTTTACACCTCAT	900
Qy	301	SerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspVal	320
Db	901	AGTCTGCTGAGGAAGGTCCTGCTCCCTGAGTCTCTCTGATCGTCACCGTCAGAG	960
Qy	321	GlyThrGluLysLysSerGluValValSerProArgTyrLeuLeuValArgGlyIle	340
Db	961	GGCACAGAGAGCTCAAGTCAGAGGTGTGTCTCCCGTTTACCTGTAGTAGGAATC	1020
Qy	341	SerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLysThr	360
Db	1021	TCCGGGGAACAAAGAAATCCACTTGTCTTGGAGCGGGATTGGTGAAGCATCAGA	1080
Qy	361	GlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla	380
Db	1081	CAAGGTTGCTGGGATCATGAACACCGTGAGCTGCTCGACCAAGTCCAGGTCGCG	1140
Qy	381	ValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla	400
Db	1141	GTGGGCTCTCTCATCTGCGTGGCCCTCGAGCTCGAGACGCTGGTGGGGAGAGCG	1200
Qy	401	ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro	420
Db	1201	CCCTTCAACCAACGCTCACAGGCTCGACGCGCTTTGTGTTTCATAGCTCACCC	1260
Qy	421	ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys	440
Db	1261	CGAGGCTGCTCGGCGCTGCTCTCAATCTGGAGGAAGAGTTCCTCTGAAGCGCTT	1320
Qy	441	ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet	460
Db	1321	CGTATGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCATGATGTCAGCACCT	1380
Qy	461	ValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAenIleLeuLeuPro	480
Db	1381	GTTCAAGGATCTCGGGAGTCTGAGCTCCGTCGCTGTTTTCATGAACATCTCTCT	1440
Qy	481	AspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla	500
Db	1441	GACAGCCACTGTGAGGAGTACTACACTTCTCCACTCAGTCTCCAGGAGTCTCTG	1500
Qy	501	AlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrVal	520
Db	1501	GCCTTGCTACTACGTGTAGAGGCGCTGGAATCGAGCCAGCTCTCTGCCCTCTG	1560
Qy	521	GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTip	540
Db	1561	GAGAGCAAGAGGTCCATGGAGCTTAAACAGGAGGCTTCCATATCCACTCGCTT	1620
Qy	541	MetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeu	560
Db	1621	ATGAAGCGTTTCTGTTTGGCCTCGTGAGCGAAGAGCTAAGGAGGCGCACTGGAG	1680

Qy 561 LeuGlyCysProValProLeuGlyVallysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGCACATGGGTCTCTCTGTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrTrpGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTACAGCAGCCTAATGCCACACCCAGAGACACCTCTGGACGCCCTCCACATGCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTGTTCGCTTGGCATTTAAACAGCTTCCAAGAGTGTGGCTT 1860
Qy 621 ProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CGGATTAACAGAACCTGGACTTGATAGCATCTTCTCTGCCCTCCAGCACTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspVallysGlyIlePheProArgAspGluSerAlaGluAla 660
Db 1921 TTGCGGAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 1981 TGTCTGTGTCCCTCTATGATGCGGATAGACCCCTCATTTGAGGACAGTGGGAAGAT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 2041 TTCTGCTCCATGCTTGGCACCCACCCACACCTGGGGCAGCTGGACCTGGGCAGCAGCATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 2101 CTGACAGACGGGGCATGAAGACCCCTGTGTGCCAAGCTGAGGCATCCCACTGCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGCACCTCTGAGAAATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACTGGGGAGCACCCACCTGTAAGGAGAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArg 780
Db 2281 GATGTAAGGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTGTTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysCysGlyLeuThrHisAlaCysTyrIleLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGGATTGCTGTGGATTGACCCATGCCCTGTACCTGAAAGATCTCCCAAAATCCTTACGACC 2400
Qy 801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCACGCTGAATCTCTGAGCTTGCAGGAAACAGGTGCACAGACCAGGGAGTAATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTCAGTGTGCTTCCAGAGTCTCCAGTGGCGCTGCAGAAAGCTGATATCTGGAGC 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGCCACGGGTTGCCAGAGTCTGGCCCTCAGCCCTCGTCAGAACCGGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCys 880
Db 2581 TTGACACACCTGTGGCTATCCACACAGCCTGGGGACAGAGGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGTATGCTGAATCAGTGCACCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACACGGCTGGCTGGGTTTCTTGTGCATGCGCTTATGGGTAACTCATGTGCTGACGCCAC 2760

Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyVallysLeuLeuCysGluValMet 940
Db 2761 CTGAGCCTTAGCATGAACCCCTGTGGAAGACAAATGCGGTGAAGCTTCTGTGCGAGGTCAATG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuLeuVallysCysHisLeuThrAlaAla 960
Db 2821 AGAAGAACCATCTTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCGCGCG 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGTCTGTGAGAGTCTGTCTCTGTGTGATCTCGAGGAGACACACCTGGAAGAGCCTGGATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAAATGCCCTGGGTGACGGTGGGGTTGTGCGCTGTGCGAGGACCTGAACCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGlu 1020
Db 3001 AACAGTGTCTGACGAGACTCGGTTTGAAGGCATGTGAGCTGACTTCTCTGATTGCTGTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCATCTCTCTTGGCCCTTCTCTGCACCCGGCATCTGACCACTCTAAACCTGTGTCAGAAAT 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTTTCGGSCCTTTGCTCTGCCACGCTCTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 3181 TTACAGATAAATTTGGGCTGTGGAAATGGCAGTAGTCCCTGTGCAAAATGAAGAGCTCTCGAG 3240
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAGGTGAGCTACTCAAGCCCCGAGTGTAAATGACGGTAGTGTGGCATTTCTTTTGATGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTTCGGCTCCCTGAAAGCCGGGCATGSCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysLysArgValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTGTGTGGGGATGAACCCAGAGCAGAGAAGCGTGTGTCGCTTCTGGCTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyGluSer 1160
Db 3421 TTCAAGAGCAGTACACGATTTGCCAAGTCTCTCTGCTGCGCCACGGCAATGGTGGTCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGAGTTGACAAACGTGGAGCAGAGCTCCCCCGCNAACCATGGCAGGCACGGAACACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTyrSerGlyValaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTAGTGTGGATATTCGGAGCCTGGTCTGAAAATGCTGAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyGlyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGGCTTGATCCCAACAGTGTGATCATGACGAGGATATGGCTTGTCTCATTAGGGAGA 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAAACAGTGTCTGATGATACCACAGCGGTGTGCTCTGGT 3720
Qy 1241 HisTrpGluArgLeuGlySerArgGlyTyrCysLeuAsnSerAlaAspAspHisSerGly 1260
Db 3721 CACTGGAGGCGGTGGGCTCTAGGGGCTGGTGTCTTAAACAGTGTGATGACCACAGCGGT 3780
Qy 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAsp 1280
Db 3781 GTGTCTTGGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGAC 3840
Qy 1281 HisSerGlyValAlaTrpSerLeuGlyAlaAlaGlyLeuGlyLeuValSerAsnSer 1300

QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 1894 TCTCTGTTGGGTGAGCAGCAGTAATGCCACCAACCCAGGAGACACCCCTGGACGCTTCCAC 1953
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 1954 TGTCTTTTCGAGACTCAAGACAAGAGTTTGTCTTCCTTGGCATATAACAGCTTCCAGAA 2013
QY 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 2014 GTGTGGCTTCGGATTAACAGAACCTGGACTGTAGTAGCATCTTCTCTCGCTCCAGAC 2073
QY 638 CysProTrpLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB 2074 TGTCCGTATTTTCGGGAATAATTCGGGTGGATGTCAAAAGGGATCTTCCCAAGAGATGAGTCC 2133
QY 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB 2134 GCTGAGGATGTCCTGTGTCTCTATGGATGCGGGATAAGACCTCATTTGAGAGCAG 2193
QY 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 2194 TGGGAAGATTCTGCTCCATGCTGGCACCCACCCACACCTGGCGACCTGGACCTGGGC 2253
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 2254 AGCAGCATCTGACAGAGCGGGCCATGAAGCCCTGTGTGCCAAGCTGAGGCATCCCCACC 2313
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 2314 TGCAGATATACAGACCTCATGTTTAGNAATGCACAGATTACCCCTGGTGTGCACGACCTC 2373
QY 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
DB 2374 TGGGAATTCGTATGCGCCACCGTAACCTAAGATCCCTCAACTGGGAGGCCACCCACCTG 2433
QY 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
DB 2434 AAGGAAGAGGATGTAAGATGCGGTGTGAAGCCTTTAAACACCCCAAAATGTTGTGGAG 2493
QY 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIle 797
DB 2494 TCTTTGAGCGTGATGTTGCTGTGATTTGACCCATGCTGTATACCTGAAGATCTCCCAATC 2553
QY 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
DB 2554 CTTAGACCTCCCCAGCGCTGAAATCTCTGAGCCTGGCAGGAAACAGGTGACAGACCG 2613
QY 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
DB 2614 GGAGTAATGCCTCTCAGTGTATGCTTTGAGAGTCTCCAGTGGCGCTGCAGAACTGATA 2673
QY 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
DB 2674 CTGGAGGACTGTGGGCATCACAGCCACCGGTGGCCAGAGTCTGGCCCTCAGCCCTCGTCAGC 2733
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGAGCTTGACACACTGTGCTCTATCCAAACACGCTGGGGAAACGAAGGTGTAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTGATCCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGATCTGAAATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCCACCTGGACACCGCTGGCTGTGGTTTCTTGTGCACCTTGGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTGACGACCTTGAGCCTTAGCATGAACCTCTGGGAAGACAATGGCGTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957

DB 2974 GAGGTCAATGAGAGAACCATCTTGTCACTCTCAGGACCTGGAGTTGGTAAAGTGTCACTC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
DB 3034 ACCCCCGGTGCTGTGAGAGTCTGTCTCTGTGTATCTCGAGGAGCAGACACCTGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCAGGACATGCCCTGGGTGACGGTGGGTGCTGCTCACTGTGCGAGGAGCTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGAT 3213
QY 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
DB 3214 TGTGTGAGGACACTCTCTTGGCCCTTCTCTGCAACCGGATCTGACCACTTAACACTG 3273
QY 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
DB 3274 GTGCAGAAATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTCGGCTTTGCCCTGTCCC 3333
QY 1058 ThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLys 1077
DB 3334 ACGTCTAACTTACAGATAATTGGGCTGTGGAAATGGCAGTACCCTGTGCAAAATAAGGAAG 3393
QY 1078 LeuLeuGluValGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1097
DB 3394 CTGTGAGGAAGTGCAGCTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTTGGCATTTCT 3453
QY 1098 PheAspGluAspAspArg 1103
DB 3454 TTTGATGAAGATGACCGG 3471
RESULT 4
US-10-216-645-1
; Sequence 1, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA
; APPLICANT: PETERS-KOTTIG, MICHAELE
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN WATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3926
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-216-645-1
Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 15 Gaps: 0
US-10-066-521-6 (1-1344) x US-10-216-645-1 (1-3926)
QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
DB 454 ATTTCAAGCTATGGAAACAAGAGGTGGCCACAGCAGCAGACAGACAGAAACAAGAAATT 513
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137

|||||
514 TCACAGCTATGGAAACAGAGGTGCCACACAGCAGCAGACAGAAACAGAGCATGGA 573
QY GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
DB GGTGACACATGGGACTACAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGTGA 633
QY ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
DB CGTCGTAGTTTTGAAACACATGCTGCTACCTGGCCGGAATAAGCAAAACGTTGGCTGGCT 693
QY PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
DB TTTGATTCAGACCGGTGGGCTTCGGCTCGCACGGTGGTTCTGCACGGAAAGTCAGGA 753
QY IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyr 217
DB ATTGGGAAATCGGCTTAGCCAGCAAGGATCGTGTGTCTGGCGGCAAGGTGGACTTAC 813
QY GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
DB CAGGGAATGTCTCTACGTCTTCTTCTCCCGTTAGAGAGATGCACGCGGAAGAAGGAG 873
QY SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
DB AGCAGTGTCAACAGATTCTCTCCAGGAGTGGCCAGACTCCAGAGGTCCGCTCACCGGAG 933
QY IleMetSerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspAspLeuGlySer 277
DB ATCATGTCCCGACCAAGAGCTGTGTTCATCATATGACGGTTTCGATGACCTGGGCTCT 993
QY ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThr 297
DB GTCTCTCAACAATATGACAAAGCTCTGCAAAAGACTGGGCTGAGAAGCAGCTCCGTTCACC 1053
QY LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB CTCATACGCACTGCTGAGGAAGTCTGCTCCCTGAGTCTTCTTGATGTCACCGTC 1113
QY ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
DB AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGTCTGTGTCTCCCGCTTACCTGTAGTT 1173
QY ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
DB AGAGAATCTCCGGGAAACAAGAATCCACTGTCTTGTGAGCGCGGATTTGGTGAGCAT 1233
QY GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
DB CAGAAGACACAAGGTTGCGTGGCATCATGAACCAACCGTGAGCTGCTCGACCACTGCCAG 1293
QY ValProAlaValGlySerLeuIleCysValAlaIleGlnLeuGlnAspValValGlyGlu 397
DB GTGCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAG 1353
QY SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
DB AGCGTGCGCCCTTCAACCAAAACGCTCACAGGCTGCACGCGCTTTGTGTTCATCAAG 1413
QY LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLys 437
DB CTCACCCCTCCAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAGAAGTTGTCTCAAG 1473
QY ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
DB CGCTTCTCCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTTGACGCTGAC 1533
QY AspLeuMetValGlnGlyLeuGlyCysLeuSerGluArgAlaLeuPheHisMetAsnIle 477
DB GACCTCATGGTTCAAGGACTCGGGAGTCTGAGCTCCGCTCTGTCTTTCATGAACATC 1593
QY LeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAsp 497

DB CTTCTCCACAGCAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGCTCTCCAGGAC 1653
QY PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
DB TTTCTGTGCCCTTGTACTACGTGTAGAGGGCTCGAAATCGAGGCAGCTCTCTGCCCC 1713
QY LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
DB CTGTACGTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGACGCTTCCATATCCAC 1773
QY SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
DB TCGCTTTGGATGAAGCGTTTCTTTGCTTGGCTCTGTGAGCGAAGACGTAAGAGGACCATG 1833
QY GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
DB GAGGTCTGTGTGGCTGTCTCCGTTCCTCTGGGGTGAAGCAGAGCTTCTGCACTGGGTC 1893
QY SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB TCTCTGTGGGTGAGCAGCCTAATGCCACCCAGGAGACACCTTGGACGCTTCCAC 1953
QY CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB TGTCTTTCCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAAACACCTTCCAGAA 2013
QY ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB GTGTGGCTTCGGATTAAACAGAACCTGGACTTGATAGCATTCTTCTCTGCTCCAGCAC 2073
QY CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB TGTCCGTATTTCGGAAAAATTCGGGTGGATGTCAAAAGGGATCTTCCAAAGATAGTCC 2133
QY AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB GCTGAGGCATGTCTGTGTGTCCTCTATGATGCGGGATAAGACCTCATTGAGGAGCAG 2193
QY TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB TGGGAAGATTCTGCTCCATGCTTGGCACCCACCCACACCTGCGGACGCTGGACCTGGGC 2253
QY SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB AGCAGCATCTGACAGAGCGGCCCATGAGACCTGTGTGCCAAGCTGAGGCACTCCACCC 2313
QY CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB TGCAAGATACAGACCTGATGTTTAGAAATGCACAGATTACCCCTGCTGTGCAGCACCTC 2373
QY TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
DB TGAGAAATCGTCTATGSCCAACCGTAACCTAAGATCCCTCAACTTGGGAGGACCCACCTG 2433
QY LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
DB AAGAAGAGATGTAGGATGGCGTGTGAAGCTTAAACACCCCAAAATGTTTGTGGAG 2493
QY SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
DB TCITTTGAGGCTGGATTGCTGTGGATTGACCCATGCTGTTTACCTGAAGATCTCCCAAATC 2553
QY LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
DB CTTACGACCTCCCGCCAGCTGAAATCTCTGAGCCTGGCAGGAAACAAAGGTGACAGACCAG 2613
QY GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
DB GGAAGTAATGCTCTCAGTGATGCTTGAGAGTCTCCAGTGCAGCTCCGCTGACAGAGCTGATA 2673
QY LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
DB CTGGAGGACTGTGGCATCACAGCACGGGTTGCAGAGTCTGGCCTTGCAGCCCTCGTCAGC 2733


```
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGGAGCTTGACACACCTGTGCCTATCCAAACACAGCTGGGGACGAAGGTGAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTGCGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCCACCTGGACACGGCTGGCTGTGGTGTCTTTCGACCTGGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTGACGCACTTGAGCTTAGCATGAACCTGTGGGAAGACATGGCGTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
DB 2974 GAGGTTCATGAGAGAACCATCTTGTCTATCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
DB 3034 ACCGCGCGGTGTGTGAGAGTCTGTCTCTGTGTGTCTCGAGAGAGACACCTGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCAGGACAATGCTGGGTGACGTTGGGTGTGCTGCATCTGTGCGAGGAGCTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGCACTTCTGAT 3213
QY 1018 CysGlyGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
DB 3214 TGCTGTGAGGCACTCTCTCTGGCCCTTTCCTGCAACCGGCATCTGACCAGTCTAAACCTG 3273
QY 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
DB 3274 GTGCAAGATACTTCAGTCCCAAGAGATGATGAAGCTGTGTTCGGCCCTTTGGCTGTCCC 3333
QY 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTrpValGlnIleArgLys 1077
DB 3334 ACCTCTAACTTACAGATAATTTGGGCTGTGGAATGGCAGTACCCCTGTGCAAAATAAGGAAG 3393
QY 1078 LeuLeuGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSer 1097
DB 3394 CTGCTGGAGGAAGTGCAGCTACTCAAGCCCGGAGTGTGTAATTGACGGTAGTTGGCATTTCT 3453
QY 1098 PheAspGluAspAspArg 1103
DB 3454 TTTGATGAAGATGACCGG 3471
RESULT 5
US-10-216-645-3
; Sequence 3, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA
; APPLICANT: PETERS-KOTTIG, MICHAEL
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN WATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3830
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-10-216-645-3
Alignment Scores:
Pred. No.: 0 Length: 3830
Score: 988.00 Matches: 988
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.51% Indels: 0
DB: 15 Gaps: 0
US-10-066-521-6 (1-1344) x US-10-216-645-3 (1-3830)
QY 116 GluIleSerGlnAlaMetGluGlnGluGlyValAlaThrAlaAlaGluThrGluGluGlnGly 135
DB 451 GAAATTTTCAACAGCTATGGAACAAGAGGTGCCACAGCAGACAGACAGAGAACAAGGA 510
QY 136 HisGlyCysAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
DB 511 CATGGAGGTGACACATGGGACTACAAGAGTCAAGTGTATGACCAAAATTCGCTGAGGAGAG 570
QY 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
DB 571 GATGTAGTCTGTAGTTTTTGAACAACATGCTGCTGACTGGCCGGAAATGCAAAACGTTGGCT 630
QY 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
DB 631 GGTGCTTTTGAATTCAGACCGGTGGGGCTTCGGGCTCGCACGGTGTCTTGCACGGAAG 690
QY 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
DB 691 TCAGGAATTTGGGAAATCGGCTCTAGCCAGAGGATCGTGTGTCTGGGCGCAAGGTGGA 750
QY 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
DB 751 CTCCTACAGGGAATGTTCTCTCTAGCTCTTCTCTCCCGTTAGAGAGATGCACGCGAAG 810
QY 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
DB 811 AAGGAGAGCAGTGTACAGAGATTTCATCTCCAGGGAGTGGCAGACTCCAGGCTCCGGTG 870
QY 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeu 275
DB 871 ACGGAGATCATGTCCCAACAGAAAGCTGTGTTCATCATTTGACGGTTCGATGACCTG 930
QY 276 GlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPro 295
DB 931 GGCTCTGTCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGCAGCCTCCG 990
QY 296 PheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
DB 991 TTCACCTCATACGACAGTCTGCTGAGGAAGTCTCTCTCTGAGTCTCTCTCATCGTC 1050
QY 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
DB 1051 ACCGTGAGAGACGTGGGCACAGAGAAGCTCAAGTTCAGAGGTGCTGTCTCTCCCGTTACCTG 1110
QY 336 LeuValArgGlyIleSerGlyGluGlnArgTleHisLeuLeuLeuGluArgGlyIleGly 355
DB 1111 TTAGTTAGAGGAATCTCCGGGGAACAAGAAATTCACATTGCTCTTGAAGCGCGGATTTGGT 1170
QY 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
DB 1171 GAGCATCAGAAGACACAAGGGTTGCGTGCATCATGAACACCGTAGCTGTCTCGACAG 1230
QY 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
DB 1231 TGCCAGGTGCCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGTCAGCTGTCAGGAGCTGGTG 1290
QY 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
DB 1291 GGGGAGAGCGTGGCCCCCTTCAACCAACCGCTCAAGGCGCTGCACGCGCTTTTGTGT 1350
```

QY 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValVal 435
DB 1351 CATAGCTTACCCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGGAGAAAGAGTGTCT 1410
QY 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
DB 1411 CTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGATAGGAAGTCAAGTGTTCAC 1470
QY 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
DB 1471 GGTGACGACCTCATGGTTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTCTTTTACATG 1530
QY 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTrpPhePheHisLeuSerLeu 495
DB 1531 AACATCTCTTCCAGACGACCTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTC 1590
QY 496 GlnAspPheCysAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeu 515
DB 1591 CAGGACTTCTGTGCCCTTGTACTACGTGTAGAGGCTTGGAAATCGAGCCAGCTCTC 1650
QY 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
DB 1651 TGCCCTCTGTAGTTCAGAACACAAAGAGTCCATGGAGCTTAAACAGGACAGGCTTCCAT 1710
QY 536 IleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
DB 1711 ATCCACTCGCTTGGATGAGCGGTTCGTGTGGCCCTCGTGAGCGAAGAGCTAAGGAGG 1770
QY 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
DB 1771 CCAGTGGAGGTCCTGCTGGGCTGCCCGTTCCTCCCTGGGGGTGAAGAGAGTTCCTGAC 1830
QY 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
DB 1831 TGGGTCTCTGTGGGTGACAGCTAAATGCCACCCAGGAGACACCCCTGGAGGCC 1890
QY 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
DB 1891 TTCACCTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACAGCTTC 1950
QY 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
DB 1951 CAAGAAGTGTGGCTCCGATTAACAGAACTCGACTTGATAGCATCTTCTCTCTGCTC 2010
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
DB 2011 CAGCACTGTCCGTAATTCGGGAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGAT 2070
QY 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675
DB 2071 GAGTCCGCTGAGGCATGTCTGTGCTCTATGGATCGGGATAGACCCCTCATTTGAG 2130
QY 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
DB 2131 GAGCAGTGGGAAGATTTCTGCTCCATGCTTGGCACCACCCACACCTGGGGCAGCTGGAC 2190
QY 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
DB 2191 CTGGGACAGCAGCATCTGACAGCGGGCCATGAAGACCCCTGTGTGCCAAGCTGAGGCAT 2250
QY 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
DB 2251 CCCACCTGCAGATACAGACCTCATGTTTAGAATGCACAGATTACCCCTGGTGTGAG 2310
QY 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThr 755
DB 2311 CACCTCTGAGAAATCGTCATGGCCAAACCGTAACCTAAGATCCCTCAACTGGGGAGGCACC 2370
QY 756 HisLeuLysGluAlaAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
DB 2371 CACCTGAAGGAGAGGATGTGAAGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTTG 2430
QY 776 LeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795

DB 2431 TTGGAGTCTTTGAGGCTGGATTGCTGTGGATTGACCCATGCCCTGTACCTGAAGATCTCC 2490
QY 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThr 815
DB 2491 CAAATCTTTACGACCTCCCCAGCCTGAAATCTCTGAGCCCTGGCAGAAACAAGGTGACA 2550
QY 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
DB 2551 GACCAGGAGTAATGCTCTCAGTGATGCTTCCAGAGTCTCCAGTGCGCCCTGCAGAG 2610
QY 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
DB 2611 CTGATCTGGAGACTGTGGCATCACAGCCACGGGTTCACAGAGTCTGGCCCTCAGCCCTC 2670
QY 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
DB 2671 GTCCAGCAACCGAGCTTGACACACCTGTGCTATCCAAACACAGCCTGGGGACGAAGGT 2730
QY 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
DB 2731 GTAAATCTACTGTGCTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTG 2790
QY 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915
DB 2791 AATCAGTGCACCTGACACGCTGGCTGGTGTGTTTCTTGCACTTGGCTTATGGGTAAAC 2850
QY 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
DB 2851 TCATGGCTGACGACCTGAGCCTTAGCATGAACCCCTGTGGAACAACATGGCGTGAAGCTT 2910
QY 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
DB 2911 CTGTGCGAGTCTATGAGAGAACCATCTGTCTCATCTCCAGACCTGGAGTTGGTAAAGTGT 2970
QY 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975
DB 2971 CATCTCACCCCGCTGCTGTGAGAGTCTGCTCTGTGATCTCGAGGACACACACCTG 3030
QY 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995
DB 3031 AAGAGCTGGATCTCACGGACAAATGCCCTGGGTGACGGTGGGTGCTGCTGCTGCGAG 3090
QY 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
DB 3091 GGAAGTGAAGCAAAAGAACAGTGTCTGACAGACTCGGGTTGAAGGCATGTGGACTGACT 3150
QY 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
DB 3151 TCTGATGTGTGAGGACCTCTCTTGGCCCTTCTGCAACCGGCATCTGACACAGTCTA 3210
QY 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSeralPheAla 1055
DB 3211 AACCTGGTGAGAAATAACTTCACTCCCAAGGAATGATGAAGCTGTGTTTCGGCTTTGCC 3270
QY 1056 CysProThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
DB 3271 TGTGCCACGCTCAACTTACAGATAAATGGGCTGTGAAATGGCAGTACCCCTGTGCAATA 3330
QY 1076 ArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrp 1095
DB 3331 AGGAAGCTGCTGAGGAGAGTGCAGCTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTTGG 3390
QY 1096 HisSerPheAspGluAspArgArg 1103
DB 3391 CATCTTTTGTGATGAAGATGACCGG 3414

RESULT 6

US-10-092-900A-347
; Sequence 347, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara

```

; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Rameeh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 347
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)..(3168)
US-10-092-900A-347

Alignment Scores:
Pred. No.: 0
Score: 968.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 72.02%
DB: 13
Length: 3226
Matches: 968
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

```

US-10-066-521-6 (1-1344) x US-10-092-900A-347 (1-3226)
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155
Db 249 CATGGAGTGCACATGGGACTACAGAGTCACGTGATGACCAATTCGTGAGGAGGAG 308
Qy 156 AspValArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 309 GATGTACGTGCTAGTTTGAATAACACTGCTGCTGACTGGCGGAAATGCAAACTGGCT 368
Qy 176 GlyAlaPheAspSerAspArgTTPGlyPheArgProArgThrValValLeuHisGly 195
Db 369 GGTGCTTTTGAATTCAGACCGGTGGGGCTTCGGGCTCCGACGGTGGTTCGCGGAAAG 428
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 215
Db 429 TCAGGAATTCGGAAATCGGCTCTAGCCAGGAAGATCGTGTGCTGGGCGCAAGGTGGA 488
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 235
Db 489 CTCTACCCAGGGAATGTTCTCTACGTCTTCTCTCCCGTTAGAGAGATGCAGCGAAG 548
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTTPProAspSerGlnAlaProVal 255
Db 549 AAGGAGAGCAGTGTCAAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTG 608
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeu 275
Db 609 ACGGAGATCATGTCCCGACCAAGAAAGGCTGTGTTTCATCATTTGACGGTTTCGATGACTG 668
Qy 276 GlySerValLeuAenAenAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPro 295
Db 669 GGCTCTGTCTCAACAATGATACAAAGCTCTGCAAGAGCTGGGCTGAGAAAGCAGCTCCG 728
Qy 296 PheThrLeuIleArgSerLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
Db 729 TTCACCTCATACGAGTCTGCTGAGGAAGTCTCTCTCTCTGAGTCTCTCTGATCGTG 788
Qy 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
Db 789 ACCGTACAGACGCTGGGCACAGAGAGTCTCAAGTCAGAGGTCGTGTCTCCCGTTACTG 848
Qy 336 LeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGly 355
Db 849 TTAGTTAGAGGAATCTCCGGGAAACAAAGAAATCCACTTGCTCTTGAGCGCGGATGGT 908
Qy 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
Db 909 GAGCATCAGAAAGACACAAAGGTTGCGTGCATCATGAACCAACCGTGAGCTGCGACGAG 968
Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
Db 969 TGCAGGTGCCCGCGGTGGGCTCTCTCATCTGCTGGTGGCCCTGCAGCTGCGAGGAGTGTG 1028
Qy 396 GlyGluSerValAlaProPheAenGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
Db 1029 GGGGAGAGCGTCCGCCCTTCAACCAAAACGCTCACAGGCTGCACGCGCTTTGTGTGT 1088
Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAenLeuGluGluArgValVal 435
Db 1089 CATCAGCTCACCCCTCGAGGCGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTC 1148
Qy 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAenArgLysSerValPheAsp 455
Db 1149 CTGAAGCGCTTTCGCCGTATGGCTGTGGAGGAGTGTGGATAGGAAGTCAAGTGTGTTGAC 1208
Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGlnLeuArgAlaLeuPheHisMet 475
Db 1209 GGTGACGACCTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATG 1268
Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeu 495
Db 1269 AACATCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTC 1328

```


; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 23

; LENGTH: 3900

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(3600)

US-10-677-943-23

Alignment Scores:

Pred. No.: 0 Length: 3900

Score: 616.00 Matches: 1002

Percent Similarity: 99.21% Conservative: 0

Best Local Similarity: 99.21% Mismatches: 4

Query Match: 45.83% Indels: 8

DB: 12 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-677-943-23 (1-3900)

Qy	98	IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGluIle	117
Db	568	ATTTCACAGCTATGGAACAGAGGTGCCACAGCAGCAGAGACAGAGAACAGAAATT	627
Qy	118	SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly	137
Db	628	TCACAAGCTATGGAACAGAGGTGCCACAGCAGCAGAGACAGAGAACAGGACATGGA	687
Qy	138	GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal	157
Db	688	GGTGACACATGGGACTACAGAGTACCGTGTATGACCAAAATTCGCTGAGGAGGAGATGTA	747
Qy	158	ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla	177
Db	748	CGTCGTAGTTTGAANAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	807
Qy	178	PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly	197
Db	808	TTTGATTTCAGACCGGTGGGGCTTCGGGCTCGACGGTGGTTCGACGGAAAGTCAGGA	867
Qy	198	IleGlyLysSerAlaLeuAlaArgGlyIleValLeuCysTrpAlaGlnGlnGlyLeuTyr	217
Db	868	ATTGGGAATCGGCTCTAGCCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	927
Qy	218	GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu	237
Db	928	CAGGGAATGTTCTCTAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	987
Qy	238	SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu	257
Db	988	AGCAGTGTACACAGAGTTCATCTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTACG	1047
Qy	258	IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySer	277
Db	1048	ATCATGTCCCGACCAAGAGGCTGTGTTCATCATGTGACGGTTTCGATGACCTGGGCT	1107
Qy	278	ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThr	297
Db	1108	GTCCTCAACATGACACAAAGCTCTGCAAGAGTGGGCTGAGAGCAGCCTCCGCTTACC	1167
Qy	298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	1168	CTCATACGCGTCTGCTGAGGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1227
Qy	318	ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal	337
Db	1228	AGAGACGTGGGCACAGAGAGTCAAGTCAGAGGTCTGCTCTCCCGCTTACCTGTAGTT	1287
Qy	338	ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis	357
Db	1288	AGAGGAATCTCCGGGGAACAAGAAATCCACTTGTCTTGTAGCGCGGGAATGGTGAGAT	1347
Qy	358	GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysGly	377

Db	1348	CAGAAACACAAAGGGTTGCGTGGCATCAT-CAACAACCGTGAGCTGCTCGACCAAGTCCCA	1406
Qy	377	nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGly	397
Db	1407	GGTGCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGCGAGACGTGGTGGGGGA	1466
Qy	397	uSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG	417
Db	1467	GAGCGTCGCCCTTCAACCAACGCTCACAGGCTGCACGCGCTTTTGC-GTTTCATC	1525
Qy	417	lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL	437
Db	1526	AGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTCGAGGAAGAGTTGTCTCTGA	1585
Qy	437	ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA	457
Db	1586	AGCGCTTCTCCGTATGGCTGTGAGGAGGTGGAAATAGAAATCAGTGTGTGATGGTG	1645
Qy	457	spAspLeuMetValGlnGlnGlyLeuGlySerGluLeuArgAlaLeuPheHisMetAsnI	477
Db	1646	ACGACCTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTTCACATGAACA	1705
Qy	477	leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA	497
Db	1706	TCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTCCAGG	1765
Qy	497	spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP	517
Db	1766	ACTTCTGTGCGCGCTGTACTAGCTGTGAGGCGCTGGAAATCGAGCCAGCTCTCTGCCC	1825
Qy	517	roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH	537
Db	1826	CTCTGTACGTTGAGACACAAAGAGTCCATGAGGCTTTAAACAGGCAAGCTTCCATATCC	1885
Qy	537	isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL	557
Db	1886	ACTCGCTTTCGATGAGCGTTCCTTGTGTCCTCGTGAGCGAAGAGCTAAGGAGGCCAC	1945
Qy	557	euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV	577
Db	1946	TGGAGGTCTCTGCTGGGCTGTCCCGTTCCTCTGGGGGTGAAGCAGAGCTTCTGCACTGG	2005
Qy	577	alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH	597
Db	2006	TCTCTCTGTTGGGTACAGACGCTAATGCCACACCCCGAGGAGACACCTCGACGCGCTTC	2065
Qy	597	isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG	617
Db	2066	ACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACAGCTTCCAAG	2125
Qy	617	luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH	637
Db	2126	AAGTGTGGCTTCGATTAACAGAACCTGGACTGTGATAGCATCTTCTTCTGCTCCGAGC	2185
Qy	637	isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS	657
Db	2186	ACTGTCCGTATTTGGGGAATAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGAGATGAGT	2245
Qy	657	erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlu	677
Db	2246	CCGCTGAGGCATGTCTGTGTGCTCCCTCTATGATGATGCGGATAGACCTCATTTGAGGAGC	2305
Qy	677	lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG	697
Db	2306	AGTGGGAAGATTTCTGTCTCATGCTTGGCACCCACACCTCGCGCAGCTGACCTGG	2365
Qy	697	lySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT	717
Db	2366	GCAGCAGCATCTCGACAGAGCGGCCCATGAAGACCTGTGTGCCACAGCTGAGGATCCCA	2425
Qy	717	hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL	737

Db 2426 CCTGCAAGATACAGACCCCTGATGTTTACAAATGCACAGATTACCCCTGGTGTGCAACACC 2485
Qy 737 euTrpArgIleValMetAlaIenArgAenLeuArgSerLeuAenLeuGlyGlyThrHisL 757
Db 2486 TCTGGAGAATCGTCATGCGCAACCGTAACCTAAGATCCCTCAACTTGGGAGGACCCACC 2545
Qy 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db 2546 TGAAGGAAGAGATGTAAGGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTGTGG 2605
Qy 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db 2606 AGTCCTTCAGGCTGGATGCTGTGGATTGACCCATGCGCTGTACTGAAGATCTCCCAA 2665
Qy 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCCCTTACAGACCTCCCCACGCTGAAATCTCTGAGCCTGGCAGGAACAAGGTGACAGACC 2725
Qy 817 InGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGGAGTAAAC-GCCTCTCAGTGTAGCCTTGAGGCTCTCCCAAGTGGCCCTGCAGAGCTG 2784
Qy 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db 2785 ATACTGGAGGACTGTGGCATCACAGCCACGGGTTGCCAGAGCTCTGGCCTCAGCCCTCGTC 2844
Qy 857 SerAenArgSerLeuThrHisLeuCysLeuSerAenAenSerLeuGlyAsnGluGlyVal 876
Db 2845 AGCAACCGAGCTTGACACACCTGTGCTTATCCAAACACAGCCTGGGGACCAAGAGTGTA 2904
Qy 877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db 2905 AATCTACTGTGTGCATCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGATGCTGAAT 2964
Qy 897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db 2965 CAGTGGCCACCTGGACACGGCTGGCTGGG-TTCTCTTGACCTTGGCTTATGGGTAACTC 3023
Qy 916 rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAenGlyValLysLeuLe 936
Db 3024 ATGGCTGACGACCTCGACCTTAGCATGAACCCCTGTGGAAGACAATGGCGTGAAGCTTCT 3083
Qy 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi 956
Db 3084 GTGGAGGTCTATGAGAGACCATTTGTTCATCTCCAGACCTGGAGTTGGTAAAGTGTC 3143
Qy 956 sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy 976
Db 3144 TCTCACCCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAA 3203
Qy 976 sSerLeuAspLeuThrAspAenAlaLeuGlyAspGlyGlyValAlaLeuCysGluG 996
Db 3204 GAGCCTGGATCTCAGGACAAATGCCCTGGGTGAGCGTGGGTGCTCGCTGTGCGAGGG 3263
Qy 996 yLeuLysGlnLysAenSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe 1016
Db 3264 ACTGAGCAAAAGAACAGTGTTCAGCAGACTCGGGTTGAAGCATGTGAGACTGACTTC 3323
Qy 1016 rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs 1036
Db 3324 TGATTGCTGTGAGGACACTCTCTTGGCCCTTTCTGTCAACCGGCATCTGACCACTTAA 3383
Qy 1036 nLeuValGlnAenAenPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCy 1056
Db 3384 CCTGGTGAGAAATAACTTTCAGTCCCAAGGAATGATGAAGCTGTGTGCTGGCCTTTGCGCTG 3443
Qy 1056 sProThrSerAenLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleAr 1076
Db 3444 TCCCACTGTCTAACTTACAGATAATTGGCTGTGGAATGGCAGTACCTGTGCAAAATAG 3503
Qy 1076 gLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHi 1096
Db 3504 GAAGCTGTGGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTGACGCTAGTTGCA 3563

Qy 1096 sSerPheAspGluAspAspArg 1103
Db 3564 TTCTTTTGTGATGAAGATGACCGG 3585

RESULT 8

US-10-399-443-23
; Sequence 23, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; Secretary, Department of Health & Human Services, The National Institute of
; Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10399,443
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 3900
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3600)
; OTHER INFORMATION:
US-10-399-443-23

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-399-443-23 (1-3900)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 568 ATTTCAAGCTATGGAACAGAGAGGTGCCACAGCAGCAGACAGACAGAGAAAT 627
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 628 TCACAAAGCTATGGAACAGAGAGGTGCCACAGCAGCAGACAGACAGAAACAAAGGACATGGA 687
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACACATGGGACTACAAAGAGTCACTGCTGCTGACCAAAATTCGCTGAGGAGGAGTGA 747
Qy 158 ArgArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyVal 177
Db 748 CGTGTAGTATTTGAAAACACTGCTGCTGACTGCCCGAATTCGCAACGTTGGCTGGTCT 807
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 808 TTTGATTTCAGACCGGTGGGCTTCCGCCCTCGCACGCTGGTCTCGCACGGAAGTCAGGA 867
Qy 198 IleGlyLysSerAlaLeuAlaArgGlyIleValLeuCysTrpAlaGlnGlyLeuTyr 217
Db 868 ATTTGGGAAATCGGCTCTAGCCAGAGAGATCGTCTGCTGCTGGCGCGCAAGGTGACTCTAC 927
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237

Db 928 CAGGGAATGTTCTCTACGTCTTCTTCTCCCGTTAGAGAGATCAGCGGAAGGAG 987
Qy 238 SerSerValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGlu 257
Db 988 AGCAGTGTACAGAGATTCTCTCAGGGAGTGGCCAGACTCCAGGCTCCGGTACGGAG 1047
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db 1048 ATCATGTCCGACACAGAAAGCGTGTGTTCATATTGACGGTTTCGATGACTGGGCTCT 1107
Qy 278 ValLeuAsnAsnAspThrIleLeuGlyCysIleAspTyrAlaGluLeuGlnProPheThr 297
Db 1108 GTCCCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAGAACAGCCCTCCGTTACC 1167
Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1168 CTCATACGAGTCTGCTGAGGAAGTCTCTGCTCCCTGAGTCTCTCTGATCGTCAACGTC 1227
Qy 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1228 AGAGACGTGGGCACAGAGAAAGCTCAAGTCAGAGGTCTGTCTCCCGTTACTCTGTAGT 1287
Qy 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 1288 AGAGGAATCTCCGGGGAACAAGAAATCCACTTGTCTCTTGAGCGGGGATTGGTGAGCAT 1347
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysG 377
Db 1348 CAGAGACACAGAGGTTCGGTGGCATCAT-CAACAAACCGTGAGCTGCTCGAACGAGTGC 1406
Qy 377 nValProIleValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyG 397
Db 1407 GGTGCGCCGCGTGGCTCTCTCATCTGGCTGGCCCTGCAGCTGCAGGACGTGGTGGGGA 1466
Qy 397 uSerValIleProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db 1467 GAGCGTCGCCCCCTTCACACCAACGCTCACAGGCTGCAGCGCGCTTTTTCG- GTTTTCATC 1525
Qy 417 InLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db 1526 AGCTCACCCCTCGAGCGGTGGTCCGGCGTGTCTCAATCTGGAGGAAGAGATTGCTCTGA 1585
Qy 437 ysArgPheCysArgMetAlaValGluGlyValTyrProAsnArgLysSerValPheAspGly 457
Db 1586 AGCGCTTCTGCCGTATGCTGTGGAGGAGTGTGAATAGGAAGTCAGTGTTCGATGGT 1645
Qy 457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db 1646 ACGACCTCATGGTTCAAGGACTCGGGAGTCTGAGCTCCGTCTGTTCACATGAACA 1705
Qy 477 LeLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGln 497
Db 1706 TCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGG 1765
Qy 497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db 1766 ACTTCTGTGGCCCTTGTACTACGTGTAGAGGGCTCGAATAATCGACGCTCTCTGCC 1825
Qy 517 roLeuTyrValGlnLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db 1826 CTCGTGTAGTTGAGAAGCAAGAGAGTCCATGGAGCTTAAACAGGACGGCTTCCATATCC 1885
Qy 537 IsSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL 557
Db 1886 ACTCCGCTTTGGATGAAGCGTTCTTTTGGCCTCGTGAGCGAAGACGTAAAGAGGCCAC 1945
Qy 557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisIstrp 577
Db 1946 TGGAGGTCTCTGGGCTGTCCCGTTCCTTGGGGTGAACGACAGAGCTTCTGCACTGG 2005
Qy 577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db 2006 TCTCTCTGTGGGTTCAGCAGCCCTAATGCCCACACCCAGGAGACACCCCTGGACGCTTCC 2065

Qy 597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db 2066 ACTGTCTTTTCGAGACTCAAGACAAAGATTGTTTCGCTTGGCATTTAAACAGACTTCCNA 2125
Qy 617 luValTyrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db 2126 AAGTGTGGCTTCGATTAACAGAACTGGACTTGATAGCATCTTCTTCTGCTCCAGC 2185
Qy 637 isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS 657
Db 2186 ACTGTCGTATTTTCGGAATAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGAGATGAGT 2245
Qy 657 erAlaGluAlaCysProValValProLeuTyrMetArgAspLysThrLeuIleGluGluG 677
Db 2246 CCGCTGAGGACATGTCCTGTGTCTCTATGATGTCGGGTAAACACCTCATTTGAGGAGC 2305
Qy 677 InTyrGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db 2306 AGTGGGAAGATTTTCTGCTCCATGCTGGCACCCACCCACACCTCGCGCAGCTGGACCTGG 2365
Qy 697 lySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT 717
Db 2366 GCACGACATCTCTACAGAGCGGCCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCC 2425
Qy 717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db 2426 CCTGCAGATACAGACCTCTGATGTTTGTAGAAATGACAGATATACCCCTGGTGTGCAAC 2485
Qy 737 euTyrArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db 2486 TCTGGAAATCGTTCATGCCCAACCGTAACCTTAAGATCCCTCAACTTGGGAGCACCCACC 2545
Qy 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db 2546 TGAAGGAAGAGGATGTAAAGATGGCGTGTGAAGCCTTTAAACACCCCAAAATGTTTGTGG 2605
Qy 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db 2606 AGTCTTTGAGGCTGGATTGCTGTGGATTGACCCATGCTGTACTTACAGAGATCTCCCA 2665
Qy 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCCTTACGACCTCCCCAGCGCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACC 2725
Qy 817 InGlyValMet-ProtoLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGAGTAAAC-GCCTCTCAGTGATGCCCTTGAGGGTCTCCAGTGCGCCCTGCAGAAGCTG 2784
Qy 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db 2785 ATACTGGAGACTGTGGCATCACAGCCACGGGTGGCCAGAGTCTGGCCTCAGCCCTCGTC 2844
Qy 857 SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal 876
Db 2845 AGCAACCGGAGCTTGACACACCTGTGCTATCCAACAACAGCCCTGGGGAACGAAGGTGTA 2904
Qy 877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db 2905 AATCTACTGTGTGATCCATGAGGCTTCCCCACTGTGTAGTGTGACAGAGGCTGATCTGA 2964
Qy 897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db 2965 CAGTGCACCTCCGACACGGCTGGCTGTGG-TTCTCTTGACATTTGGCTTATGGSTAACTC 3023
Qy 916 rTyrLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe 936
Db 3024 ATGGCTGACGACCTGAGCCTTAGCATGAACCTGTGGAAGACAAATGGCGGTAAGCTTCT 3083
Qy 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi 956
Db 3084 GTGCGAGGTTCATGAGAGAACCCTCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTC 3143


```
OY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 10
US-10-399-443-1
; Sequence 1, Application US/10399443
; Publication NO. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR FILING DATE: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-399-443-1

Alignment Scores:
Pred. No.: 0 Length: 1157
Score: 385.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.65% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-399-443-1 (1-1157)

OY 252 GlnAlaProValThrGluLeuMetSerArgProGluArgLeuLeuPheIleAspGly 271
Db 1 CAAGCTCCGGTGACGGAGATCATGTCCTCCAGCAAGAGGCTGTTGTTTCATCATTCGCGT 60

OY 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrrAlaGlu 291
Db 61 TTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCACAAAGACTGGGCTGAG 120

OY 292 LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
Db 121 AAGCAGCTTCCTGTTACCTCATAGCAGTCTGCTGAGCAAGGCTCTGCTCCCTGAGTCC 180

OY 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
Db 181 TTCCTGATCGTCACCGTCAGAGACGTGGGCGACAGAGAAGCTCAAGTCAGAGGTCGTGTCT 240

OY 332 ProArgTrrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGlu 351
Db 241 CCCCCTTACCTGTTAGTAGAGGAATCTCCGGGGAACAAGAATCCACTTCTCTCTTGGAG 300

OY 352 ArgGlyIleGlyHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
Db 301 CCGGGGATTTGTGACATCAGAGACACAGAGGTTGCGTGGCATCATGACACACCGTGAG 360

OY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACCACTGCGAGGTCGCCCGCTGCTCTCTCATCTGCGTGGCTCGAGCTG 420

OY 392 GlnAspValValGlyLysSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
Db 421 CAGGACGTGGTGGGGAGAGCGTCCGCCCTTCAACCAAGAGCTCACAGGCTCGAGGCC 480
```

```
OY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431
Db 481 GCCTTTGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGGAG 540

OY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrrAsnArgLys 451
Db 541 GAAAGAGTTGTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGAATAGGAAG 600

OY 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
Db 601 TCAGTGTGTTGACGGTGACGACCTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCT 660

OY 472 LeuPheHisMetAenIleLeuLeuProAspSerHisCysGluGluTrrTrrPhePhe 491
Db 661 CTGTTTCACATGACATCTCTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTTCTC 720

OY 492 HisLeuSerLeuGlnAspPheCysAlaLeuTrrTrrValLeuGluGlyLeuGluIle 511
Db 721 CACCTCAGTCTCCAGGACTTCTGTGCCGCTTGTACTACGTGTAGAGGCGCTGGAATC 780

OY 512 GluProAlaLeuCysProLeuTrrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCCAGCTCTCTGCCCTCTGTACGTTGAGAAGACAAAGAGGTCATGGAGCTTAAACAG 840

OY 532 AlaGlyPheHisIleHisSerLeuTrrMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGGCTTCCATATCCACTCGCTTGGATGAAGGCTTCTTGTGGCTCTGAGGCGAA 900

OY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GACGTAAGGAGGCGCACTGGAGGCTCTGCTGGGCTGTCCGCTTCCCTGGGGTGAAGCAG 960

OY 572 LysLeuLeuHisTrrValSerLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGACCTGGGCTCTCTGTTGGGTCAGAGCTTAATGCCACACCCAGGAGAC 1020

OY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGACGCTTCCACTGCTCTTTTCGAGACTCAGACAAAGAGTTGTTCGCTTGGCA 1080

OY 612 LeuAsnSerPheGlnGluValTrrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1081 TTAACAGCTTCAAGAAAGTGTGCTTCCGATTAAACCAGAACCTGGAGCTTGATAGCATCT 1140

OY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 11
US-10-677-943-3
; Sequence 3, Application US/10677943
; Publication NO. US20040072297A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Nelson, Lawrence
; APPLICANT: Tong, Zhi-Bin
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64790
; CURRENT APPLICATION NUMBER: US/10/677,943
; CURRENT FILING DATE: 2003-10-01
; PRIOR FILING DATE: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US02/09776
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1075
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-677-943-3

Alignment Scores:
Pred. No.:      8,37e-192      Length:      1075
Score:          196.00         Matches:    252
Percent Similarity: 99.21%      Conservative: 0
Best Local Similarity: 99.21%    Mismatches: 1
Query Match:      14.58%        Indels:     2
DB:               12           Gaps:        0

US-10-066-521-6 (1-1344) x US-10-677-943-3 (1-1075)
QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
Db 2 CTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGCCCTATCCAAACACAGC 61
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAAGGTGTAATCTACTGTGTGCATCCATGAGGCTTCCCCACTGTAGTCTG 121
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCGCTGGCTGTGG-TTCTTTGCACT 180
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TCGCGCTTATGGGTAACTCATGGCTGACGCACCTGAGCATGAACCCCTGTGGAAGA 240
QY 930 pAsnGlyVallyLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAAACCACTTGTGTATCTCCAGGACCT 300
QY 950 uGluLeuVallyLeuCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTTGGTAAATGTCACTCACCGCGCGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
QY 970 rArgSerArgHisLeuSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGAAGACCTTGGATCTCAGGCAATTCAGTCCCAAGGAATGATGAAGCT 420
QY 990 lAlaAlaLeuCysGluGlyLeuGlnCysGlnSerValLeuThrArgLeuGlyLeuLy 1010
Db 421 TGCTGCGCTGTGCGAGGACTGAAAGAAAAGAACAGTGTCTGACGAGACTCGGGGTGAA 480
QY 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCAATGTGACTGACTTCTGATTTGCTGTGAGGCACTCTCCTTGGCCCTTTCTTGCAACCG 540
QY 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACCAAGCTAAACCTGGTGCAGAAATACTCAGTCCCAAGGAATGATGAAGCT 600
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
Db 601 GTGTTGCGCCCTTTCCTGTCCACGTCTAACTTACAGATAATTTGGGTGTGGAATGGCA 660
QY 1070 nTyrProValGlnIleArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValVa 1090
Db 661 GTACCCCTGTCAANTAGGAGCTGTGGAGGAAGTGCAGCTACTCAAGCCCGGAGTCTGT 720
QY 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTGGCAATCTTTTGTATGAAGATGACCCG 760

RESULT 12
US-10-399-443-3
; Sequence 3, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; Secretary, Department of Health & Human Services, The National Institute
```

```
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1075
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-399-443-3

Alignment Scores:
Pred. No.:      8,37e-192      Length:      1075
Score:          196.00         Matches:    252
Percent Similarity: 99.21%      Conservative: 0
Best Local Similarity: 99.21%    Mismatches: 1
Query Match:      14.58%        Indels:     2
DB:               13           Gaps:        0

US-10-066-521-6 (1-1344) x US-10-399-443-3 (1-1075)
QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
Db 2 CTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGCCCTATCCAAACACAGC 61
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAAGGTGTAATCTACTGTGTGCATCCATGAGGCTTCCCCACTGTAGTCTG 121
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCGCTGGCTGTGG-TTCTTTGCACT 180
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TCGCGCTTATGGGTAACTCATGGCTGACGCACCTGAGCATGAACCCCTGTGGAAGA 240
QY 930 pAsnGlyVallyLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAAACCACTTGTGTATCTCCAGGACCT 300
QY 950 uGluLeuVallyLeuCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTTGGTAAATGTCACTCACCGCGCGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
QY 970 rArgSerArgHisLeuSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGAAGACCTTGGATCTCAGGCAATTCAGTCCCAAGGAATGATGAAGCT 420
QY 990 lAlaAlaLeuCysGluGlyLeuGlnCysGlnSerValLeuThrArgLeuGlyLeuLy 1010
Db 421 TGCTGCGCTGTGCGAGGACTGAAAGAAAAGAACAGTGTCTGACGAGACTCGGGGTGAA 480
QY 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCAATGTGACTGACTTCTGATTTGCTGTGAGGCACTCTCCTTGGCCCTTTCTTGCAACCG 540
QY 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACCAAGCTAAACCTGGTGCAGAAATACTCAGTCCCAAGGAATGATGAAGCT 600
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
Db 601 GTGTTGCGCCCTTTCCTGTCCACGTCTAACTTACAGATAATTTGGGTGTGGAATGGCA 660
QY 1070 nTyrProValGlnIleArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValVa 1090
Db 661 GTACCCCTGTCAANTAGGAGCTGTGGAGGAAGTGCAGCTACTCAAGCCCGGAGTCTGT 720
QY 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTGGCAATCTTTTGTATGAAGATGACCCG 760

RESULT 12
US-10-399-443-3
; Sequence 3, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; Secretary, Department of Health & Human Services, The National Institute
```

```
Db 601 GTGTTGGCGCTTGCTGTCCACGCTTAACCTACAGATAAATTGGGCTGTGGAATGGCA 660
Qy 1070 nTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArgValVa 1090
Db 661 GTACCTGTGTGCAATAAGGAAGCTGCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGT 720
Qy 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTTGGCATCTCTTTTGATGAAGATGACCG 760

RESULT 13
US-10-027-632-258159
; Sequence 258159, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258159
; LENGTH: 2099
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258159

Alignment Scores:
Pred. No.: 4,92e-48 Length: 2099
Score: 57.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.24% Indels: 0
DB: 16 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-027-632-258159 (1-2099)
Qy 668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db 220 ATGCGGGATAGACCCCTCATTTGAGGAGCAGTGGGAGAGATTCTGCTCCATGCTTGGCACC 279
Qy 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetIys 707
Db 280 CACCCACACCTGCGGAGCTGGACCTGGGAGCAGCATCTCGACAGAGCGGGCCATGAAG 339
Qy 708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet 724
Db 340 ACCCTGTGTGCCAAGCTGAGGCATCCACCTGCAAGATACAGACCCCTGATG 390

RESULT 14
US-10-027-632-258159
; Sequence 258159, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
```

```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258159
; LENGTH: 2099
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258159

Alignment Scores:
Pred. No.: 4,92e-48 Length: 2099
Score: 57.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.24% Indels: 0
DB: 16 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-027-632-258159 (1-2099)
Qy 668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db 220 ATGCGGGATAGACCCCTCATTTGAGGAGCAGTGGGAGAGATTCTGCTCCATGCTTGGCACC 279
Qy 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetIys 707
Db 280 CACCCACACCTGCGGAGCTGGACCTGGGAGCAGCATCTCGACAGAGCGGGCCATGAAG 339
Qy 708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet 724
Db 340 ACCCTGTGTGCCAAGCTGAGGCATCCACCTGCAAGATACAGACCCCTGATG 390

RESULT 15
US-10-029-386-10957/c
; Sequence 10957, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10957
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008753.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.37
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: B1911853.1, EVALUAE 5.00e-42
; OTHER INFORMATION: SWISSPROT HIT: P33076, EVALUAE 1.00e-09
; OTHER INFORMATION: NT HIT: G114731965, EVALUAE 1.90e+00
```

US-10-029-386-10957

Alignment Scores:

Pred. No.:	0.0365	Length:	509
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	15	Gaps:	0

US-10-066-521-6 (1-1344) x US-10-029-386-10957 (1-509)

Qy 262 ProGluArgLeuPheIleAspGlyPheAsp 273

Db 178 CCCGAGCCCTCTTTTCATCATGACGGCTTCGAT 143

Search completed: July 16, 2004, 18:42:58
Job time : 1422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 08:50:58 ; Search time 12090 Seconds
(without alignments)
4818.283 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074

Sequence: 1 MEGDKSLTFSYGLQWCLYE.....DDHSGVSWSLGAGGLEGLVS 1344

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10066521/runat_13072004_122212_11081/app_query.fasta_1.1543
-DB=genEmbl -QWMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plo -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn 1 1 8159 @runat_13072004_122212_11081 -NCPU=3
-NO MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5763.5	81.5	3885	9	AY154460 Homo sapi
2	5756	81.4	3489	6	AX478549 Sequence
3	5756	81.4	3926	6	AX704821 Sequence
4	5741.5	81.2	3885	9	AY054986 Homo sapi
5	5741.5	81.2	3900	6	AX427610 Sequence
6	5724.5	80.9	5859	6	AX459881 Sequence
7	5724.5	80.9	6939	6	AX459873 Sequence
8	5724.5	80.9	6939	6	AX459891 Sequence
9	5683.5	80.3	3830	6	AX704823 Sequence
10	2814	39.8	157141	9	AC011470 Homo sapi
11	2814	39.8	193609	9	AC024580 Homo sapi
12	2695	38.1	3470	10	AY196362 Mus muscu
13	2692	38.1	3405	10	AY329487 Mus muscu
14	2688	38.0	3447	6	AX427592 Sequence
15	2688	38.0	3447	10	AF074018 Mus muscu
16	2688	38.0	3453	10	AY329486 Mus muscu
17	2688	38.0	3480	10	AY196361 Mus muscu
18	2686	38.0	3534	10	BC053384 Mus muscu
19	2686	38.0	3454	10	AY329490 Mus muscu
20	2685	38.0	3432	10	AY329484 Mus muscu
21	2567	36.3	3375	10	AY329489 Mus muscu
22	2553.5	36.1	3372	10	AY329485 Mus muscu
23	2540	35.9	3303	10	AY329488 Mus muscu
24	2533	35.8	3303	10	AY329491 Mus muscu
25	2005	28.3	1157	6	AX427588 Sequence
26	1885.5	26.7	3282	9	AY154469 Homo sapi
27	1885	26.6	193662	2	AC127352 Mus muscu
28	1732	24.5	167509	2	AC012107 Homo sapi
29	1435	20.3	3333	9	AF442488 Homo sapi
30	1435	20.3	3339	9	AY072792 Homo sapi
31	1435	20.3	3368	9	AF479747 Homo sapi
32	1434	20.3	3365	9	BC050326 Homo sapi
33	1396	19.7	3234	9	AF482706 Homo sapi
34	1356	19.2	231371	2	AC132633 Rattus no
35	1356	19.2	323826	2	AC110429 Rattus no
36	1335	18.9	1075	6	AX427590 Sequence
37	1331.5	18.8	3132	9	AY154468 Homo sapi
38	1331.5	18.8	3300	6	AX459869 Sequence
39	1320.5	18.7	3108	6	AX417214 Sequence
40	1318	18.6	1576	10	F143559S07
41	1316.5	18.6	3190	6	AX714395 Sequence
42	1316.5	18.6	3190	9	AK056688 Homo sapi
43	1315.5	18.6	3172	6	AX743691 Sequence
44	1315	18.6	3731	9	AY116204 Homo sapi
45	1306.5	18.5	3507	9	BC028069 Homo sapi

ALIGNMENTS

RESULT 1

AY154460	3885 bp	mRNA	linear	PRI 20-FEB-2003
LOCUS				
DEFINITION	Homo sapiens NALP5 (NALP5) mRNA, complete cds.			
ACCESSION	AY154460			
VERSION	AY154460.1	GI:28436363		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 3885)			
AUTHORS	Tschopp, J., Martinon, F. and Burns, K.			
TITLE	NALP5: a novel protein family involved in inflammation			
JOURNAL	Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)			
MEDLINE	2451042			
PUBMED	12563287			
REFERENCE	2 (bases 1 to 3885)			
AUTHORS	Martinon, F., Hofmann, K. and Tschopp, J.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-SEP-2002) Institute of Biochemistry, University of Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland			
FEATURES	Location/Qualifiers			
source	1..3885			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
gene	1..3885			
CDS	1..3603			
	/gene="NALP5"			
	/note="member of the NALP protein family involved in inflammation; contains NACHT, LRR and PYD domains; similar to mouse WATER protein"			
	/codon_start=1			
	/product="NALP5"			
	/protein_id="AA018156.1"			
	/db_xref="GI:28436364"			
	/translation="MKVAGLEGLGALLSAPALVTLSTGPTCSILPKNPLPPQNLSSQCIKWEDGKSTFSYGLQWCLYELDKKEFTFKELUKKSSSTTCSIFQFELENANVECLALLHEYGASLAWAYSISIFENNUKLTSEKARDDMKRHSPPEDBATMTDQGSKEKVPGISQVQDSATAAEKEQEI SQAMEQEGATAAEETEQEI SQAMEQEGA TAAETEGQGGTMDYKSHVMTKFAEEEDVRRSFENATAADPEMOTLAGAFSDRWG PRPTVYHKGSGIGKALARRIVLCWAQGLYQGMFSYVFLPVRSMORKKSSVTE FLSEWPDQAPVTEINRPERLLFIIDGFDLGSVLNNTDKLCKOWAEKQPPFTLIR SLRAKVLPSFLIIVTRDVGTEKLKSEVSPRYLLVRGSGEORIHLLERGIQEHQ KTGSLRAMNRELDDQCPVAVGSLICVALQDVVGVSVAFPNQTLTGLHAAFPVH QLTGPRVVRCLNLEERVVKRCFMAVEGVNMRKSVDFGDDLMVOGLSESLRALFHMNILLPSDHCHEYYTFPHLSIQDFCAALYVYVLEIEPALCPYVVKTKRSMELKQAFHITHLWMKRFGLVSEVDVRRPVLVLLGCPVPLGVKQKLLHWVSLGOQPNATTPG DTLDAFCLPQDQKEFVRLALNSFQSWLPIQNQLDLIASSFCLQCHPVLKIRVDV KGIFPRDESAAECPVPLMRDKTLIEQWEDFCSMGLGTHPHLQKLDLSSILITERAM KTLCAKHPTCKIKITLFRNAQITPGVHLRIVMANRNLRLNLGLGTHLKEEDVRM ACEALHKPKLLESRLDCCGLTHACYLKI SOLITSPSLKSLSLAGNKVTDQGVNPL SDALRVQCALOKLILDECGITATGCOSLASALVSNRSLTHLCLSNNSLGNQGVNLLC RSMRLPHCSLQRLMLNCHLDITACGFLALMGNSWLTHLSLSPMPVEDNGVKLLCE VMREPSCHLQDLVKCHLTAAACESLSVSRSHLKSLLDLTDLNALDGDGVAAALCEG LKQNSVHLRLGLKACGLTSDCCALSLALSCNRHLTSLNLVQNLFSPKGMKLCESAF ACPTSNLIQITGLMKWQVPIQIRKLLLEEVQLKPRVLDGSMWHSFDEDDRYWKN"			
Alignment Scores:				
Pred. No.:	0	Length:	3885	
Score:	5763.50	Matches:	1102	
Percent Similarity:	96.42%	Conservative:	2	
Best Local Similarity:	96.24%	Mismatch:	0	
Query Match:	81.47%	Indels:	41	
DB:	9	Gaps:	1	
ORIGIN				
US-10-066-521-6 (1-1344) x AY154460 (1-3885)				
Qy	1	MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu	20	
Db	154	ATGGAAGAGAGACAAATCGCTCACCTTTTCCAGCTACCGGCTGCAATGGTGTCTCTATGAG	213	

Qy	21	LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer	40
Db	214	CTAGACAAGGAAGAAATTTTCAAGCAATTTTAAAGAAAGAAATCTTTCAGAAATCG	273
Qy	41	ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu	60
Db	274	ACACATGCTCTATTTCACAGTTTGAATCGAGANTGCCAAGCTGGNATGCTTGGCACTC	333
Qy	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu	80
Db	334	CTCTTGTCATGAGTATTATGAGCATCGTGGCTGGGTACGTCCATTAGCATCTTTGAA	393
Qy	81	AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys-----	97
Db	394	AACATGAACCTGCGAACCTCTCGGAGAGCGGATGACATGAAAGACATTCACCA	453
Qy	97	-----	97
Db	454	GAAGATCCTGAAGCAAGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCAGGAATT	513
Qy	98	-----	115er 99
Db	514	TCACAAAGCTGTGCAACAAGATAGTGCACAGCTGCAGAGACAAAGAAATTTTCA	573
Qy	100	GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlnSerGln	119
Db	574	CAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAAGAAATTTTCA	633
Qy	120	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAsp	139
Db	634	GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGACAGACATGAGAGTGAC	693
Qy	140	ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg	159
Db	694	ACATGGAGCTACAAAGATGACGTGATGACCAATTCGCTGAGGAGGAGGATGACGTCT	753
Qy	160	SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp	179
Db	754	AGTTTTGAAACACTGCTGCTGACTGGCGGAAATGCAACGTTGGCTGGTCTTTGAT	813
Qy	180	SerAspArgTrpGlyPheArgProAlaGthrValValLeuHisGlyLysSerGlyIleGly	199
Db	814	TCAGACCGGTGGGGCTTCCGGCTTCGACCGTGGTCTGCACGGAAGTTCAGAAATGGG	873
Qy	200	LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly	219
Db	874	AAATCGCTCTAGCCAGAGGATCGTGTGCTGGCGCGCAAGGTGGACTCTTACCAGGA	933
Qy	220	MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer	239
Db	934	ATGTTCTCTTACCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	993
Qy	240	ValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGluIleMet	259
Db	994	GTACAGAGTTTCACTCCAGGAGTGGCCAGACTCCAGGCTCCGTGACGAGAGATCATG	1053
Qy	260	SerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspLeuGlySerValLeu	279
Db	1054	TCCCGACCAAGAGGCTGTTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGCTC	1113
Qy	280	AsnAsnAspThrLysLysCysLysAspTrpAlaGluLysGlnProPheThrLeuIle	299
Db	1114	AACAATGACAAAGAGCTCTGCAAGAGACTGGGCTGGAAGACGAGCTCGTTCACCCCTCA	1173
Qy	300	ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp	319
Db	1174	CGAGTCTGCTGAGGAGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1233
Qy	320	ValGlyThrGluLysLysSerGluValValSerProArgTyrLeuLeuValArgGly	339
Db	1234	GTGGGCACAGAGAGAGCTCAAGTCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1293

QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 1294 ATCTCGGGGAAACAAGAAATCCACTTGTCTCTGAGCGCGGATTTGGTGACATCAGAAG 1353
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
DB 1354 ACACAGGGTTTGGTGCGATCATGAACAACCGTAGCTGCTCGACCGAGTGCAGGTGCC 1413
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlySerVal 399
DB 1414 GCCGTGGGCTCTCTCATCTGCTGGCCCTGCAGCTGCAGGACGTGGTGGGGAGAGCGTC 1473
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
DB 1474 GCCCCCTTCAACCAACACCTCACAGGCTGCACGCGCTTTGTGTTCATCAGCTCACC 1533
QY 420 ProArgGlyValValArgCysLeuAsnLeuGluArgValValLeuLysArgPhe 439
DB 1534 CCTCGAGCGGTGGTCCGCGCTGCTCAATCTGGAGGAAAGAGTTGTCTGAAGCGCTTC 1593
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
DB 1594 TGGCGTATGGCTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCAGCGGTGACGACCTC 1653
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 1654 ATGGTTCAAGGACTCGGGGAGTCTGAGCTCGCTGCTCTGTTCATCATGAACATCTCTC 1713
QY 480 ProAspSerHisCysGluGluTrpThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 1714 CGAGACCGCACTGTGAGGAGTACTACACCTCTTCACACCTCAGTCTCCAGGACTCTCTGT 1773
QY 500 AlaAlaLeuTrpTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTrp 519
DB 1774 GCCGCTTGTACTAGCTGTAGAGGCGCTGGAATCGAGCGCAGCTCTCTGCGCTCTGTAC 1833
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 1834 GTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCATATCCACTCGCTT 1893
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
DB 1894 TGGATGAAGCGTTCTGTGTTGGCTCGTGAGCGAAGACGTAAAGAGGCGCCACTGGAGTC 1953
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 1954 CTGCTGGGCTGCTCCGTTCCCTCGGGGTGAAGCAGAAGCTTCTGCATGGGTCTCTCTG 2013
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 2014 TTGGGTCAAGCAGCCTTAATGCCACCAACCCAGGAGACACCTTCGACGCGCTTCCACTGTCTT 2073
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 2074 TTCGAGACTCAAGACAAAGAGTTGTTCGCTTGGCATTTAAACAGCTTCCAAGAAAGTGTGG 2133
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 2134 CTTCGATTACAGACCACTGGACTTGTATAGCATCTTCCTCTGCTCCAGCACTGTCCG 2193
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 2194 TATTTGGGAAAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGATGAGTCCGCTGAG 2253
QY 660 AlaCysProValProLeuTrpMetArgAspLysThrIleGluGlnTrpGlu 679
DB 2254 GCATGCTCTGTGGTCCCTCTATGATGCGGATAGACCCCTCATTTGAGGAGCAGTGGGAA 2313
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 2314 GATTTCTGCTCATGCTGGGACCCACCACTTCGCGGAGCTGGACCTGGGCGAGCAGC 2373
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719

DB 2374 ATCTGTGACAGAGGGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCCTGCAAG 2433
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
DB 2434 ATACAGACCTCTGATGTTTAGAAATGCACATTAACCCCTGGTGTGCAGCACCTCTGAGA 2493
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
DB 2494 ATCGTCATGCCAACCGCTAACCTTAAGATCCTCAACTTTGGAGGCACCCACCTGAAGAA 2553
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
DB 2554 GAGGATGAAGGATGGCGTGTGAAGCCTTAAAAACACCACCAAAATGTTTGTGGAGTCTTTG 2613
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIleLeuThr 799
DB 2614 AGGCTGATTTGCTGTGGATTGACCCATGCTGTACCTGAAGATCTCCCAAAATCCTTACG 2673
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
DB 2674 ACCTCCCCAGCCTGAAATCTCTGAGCCTGCGAGGAAACAAGGTGACAGACCAGGAGTA 2733
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
DB 2734 ATGCCTCTCAGTGATGCTTTGAGAGTCTCCAGTGCCTGCGCCCTGCAGAAAGCTGATAC 2793
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
DB 2794 GACTGTGGCATACAGCCAGGTTGCCAGTCTGCGCCTGAGCCCTCGTCAGCAACCGG 2853
QY 860 SerLeuThrHisLeuCysLysSerAsnAsnSerLeuGlyAsnGlyValAsnLeuLeu 879
DB 2854 AGCTTGACACACCTGTGCTTATCCAAACAACAGCCTGGGGAACGAAGGTGTAATCTACTG 2913
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
DB 2914 TGTGATTCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCAC 2973
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaMetGlyAsnSerTrpLeuThr 919
DB 2974 CTGACACGCTGGCTGTGGTTTTCTTTCACCTTGGCTTATGGGTAACTCATGGCTCAGC 3033
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
DB 3034 CACTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGCGAGGTC 3093
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuLysValLysCysHisLeuThrAla 959
DB 3094 ATGAGAAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACGCC 3153
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
DB 3154 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGGCTGGAT 3213
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGlyLeuLysGln 999
DB 3214 CTCACGGACAATGCCCTGGGTGACGTTGGGTTGCTGCACCTGTGCGAGGACCTGAAGCAA 3273
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCys 1019
DB 3274 AAGAACAGTGTCTGACGAGACTCGGTTGAAGGCATGTGGACTGTGACTTCTGATTTGCTGT 3333
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
DB 3334 GAGGCACTCTCTTGGCCCTTTCTCGCAACCGGCATCTGACCACTGTAAACCTGGTGCAG 3393
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
DB 3394 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGGCTGTCCACCTCT 3453
QY 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu 1079

```
Db 3454 AACTTACAGATAATTGGGCTGTGGAATGGCAGTACCTGTGCAATAAAGGAAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLeuLysProArgValVallileaspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAAGTGCAGCTACTCAAGCCCGAGTGTGTAATTGACGGTAGTGGCATCTTTTGAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588

RESULT 2
LOCUS AX478549 3489 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 3 from Patent WO248362.
ACCESSION AX478549
VERSION AX478549.1 GI:22217318
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ramkumar,J. and Arvizu,C.
TITLE Embryogenesis associated proteins
JOURNAL Patent: WO 0248362-A 3 20-JUN-2002;
INCYTE Genomics, Inc. (US)
FEATURES
source
1..3489
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474830CB1"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3489
Score: 5756.00 Matches: 1102
Percent Similarity: 95.34% Conservative: 2
Best Local Similarity: 95.16% Mismatches: 0
Query Match: 81.37% Indels: 54
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AX478549 (1-3489)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPhelysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTGACACAGGAAGAATTTTCAGCAATTCAGGAATTTACTAAAGAAGAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCCACAGTTTGAATCGAGAATGCCAAGCTGGAAATGCTCGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCTTGATGATATTATGGAGCATCGCTGGCCCTGGGCTACGTCATTAGCATCTTTTGA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLys----- 96
Db 241 AACATGAACCTTCGACACCTCTTCGAGAGAGCCAGGATGACATGAAATAATTCACAGAA 300
Qy 96 ----- 96
Db 301 GATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCCAGAAAAATAA 360
Qy 96 ----- 96
Db 361 TATGGCATAGCTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAAACAC 420
Qy 97 -----LysIleSerGlnAlaMetGluGlnGly 106
```

```
Db 421 AAGTATGTTGGAAATTCATTCTTTTTCGCAAAATTTTCACAGCTATGGAAACAGAGGT 480
Qy 107 AlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAlaMetGluGlnGluGlyAla 126
Db 481 GCCACAGCAGCAGACAGACAGAAAGAAATTTTCACAGCTATGGAACAAGAGGTGCC 540
Qy 127 ThrAlaAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHis 146
Db 541 ACAGCAGCAGACAGACAGAAAGAAACAGGACATGGAGGTGCACATCGGACTACAAGAGTCAC 600
Qy 147 ValMetThrLysPheAlaGluGluGluAspValArgSerPheGluAsnThrAlaAla 166
Db 601 GTGATGACCAAAATTCGTGAGGAGGAGGATGTACGTCTAGTATTTTGAACACATCGTCTGCT 660
Qy 167 AspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArg 186
Db 661 GACTGGCCGGAATGCAAAAGCTTGGCTGGTCTTTTGAATTCAGACCCGTTGGGCTTCCGG 720
Qy 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArg 206
Db 721 CCTCGCACGGTGGTCTCTGCACGGAAGTTCAGGAATTTGGGAAATCGGCTCTAGCCAGAA 780
Qy 207 IleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 781 ATCGTCTGTGCTGGGCGCAAGGTGACCTCTACAGGAATGTTCTCTACGTTCTTCTTC 840
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 841 CTCCTCGTTAGACAGATGACGCGGAAGAGGACAGAGTGTCTACAGAGTTTCATCTCCAGG 900
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 901 GAGTGGCCAGACTCCAGAGGCTCCGGTGACGGAGATCATGTGCCGACGAGAAAGGCTGTG 960
Qy 267 PheIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCys 286
Db 961 TTCTATCATTTGACGGTTCGATGACCTGGGCTCTGTCTCTCAACANTGACACAAAGCTCTGC 1020
Qy 287 LysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeuArgLysVal 306
Db 1021 AAAGACTGGGCTCAGAAAGCAGACCTCGTTCCACCTCATACGCAAGTCTGTGTGAGGAAGT 1080
Qy 307 LeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLys 326
Db 1081 CTGCTCCCTGAGTCTTCTCTGATCGTCACCGTCAGAGACGTGGGCGACAGAGAAGCTCAAG 1140
Qy 327 SerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIle 346
Db 1141 TCAGAGGTGTGTCTCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGACAAAGNATC 1200
Qy 347 HisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIle 366
Db 1201 CACTTGTCTCTTGAGCGCGGATTTGGTGAGCATCAGAAAGACAAAGGTTGCGTGCATC 1260
Qy 367 MetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCys 386
Db 1261 ATGAAACAACGTCAGTGTCTCGACCAAGTCCAGGTGCCCGCTGGGCTCTCTCATCTGC 1320
Qy 387 ValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db 1321 GTGGCCCTGACGCTGACGAGCGTGGTGGGGAGAGCGTGCCTCCCTTCAACCAACGCTC 1380
Qy 407 ThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArg 426
Db 1381 ACAGGCGCTGCACGCCGCTTTTGTGTTTCATCAGCTCACCCCTCAGGCGGTGGTCCGCGC 1440
Qy 427 CysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGly 446
Db 1441 TGTCTCAATCTGAGGAAAGAGTTGCTCTGAAGCGCTCTGCGCGTATGGCTGGGAGGGA 1500
Qy 447 ValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGlu 466
```

Db 1501 GTGTGGAAATAGGAAGTCACTGTTTCAAGCGTGACGACCTCATGGTTCAAGGACTCGGGAG 1560
Qy 467 SerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGlu 486
Db 1561 TCTGAGCTCCGTGCTCTGTTTCAATGAACATCTCTTCTCCAGACAGCCACTGTGAGGAG 1620
Qy 487 TyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaLeuTyrTyrValLeu 506
Db 1621 TACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCGCGCTGTACTAGGTGTTA 1680
Qy 507 GluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSer 526
Db 1681 GAGGCGCTGGAAATCGACGAGCTCTGTGCCCTCTGTACGTTGAGAACAAAGAGGTCC 1740
Qy 527 MetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTyrMetLysArgPheLeuPhe 546
Db 1741 ATGGAGCTTAAACAGGCAAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTGTGTT 1800
Qy 547 GlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValPro 566
Db 1801 GGCCTCGTGAGCGAAGACGTAAAGGAGGCCACTGGAGGTCTCTGTGGGCTGTCCGGTTCCC 1860
Qy 567 LeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAla 586
Db 1861 CTGGGGGTGAAGCAGAACTTCTGCACTGGGTCTCTCTGTTGGGTCAAGACCTTAATGCC 1920
Qy 587 ThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGlu 606
Db 1921 ACCACCCAGAGAGACACCTCGACGCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAG 1980
Qy 607 PheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeu 626
Db 1981 TTTGTTCCTTGGCAATTAACAGCTTCCAGAGGTGGCTTCGGATTAAACAGAACCTG 2040
Qy 627 AspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgVal 646
Db 2041 GACTTGATAGCATCTTCTCTGCTCCAGCACTGTCCGTATTTCGGGAAATTCGGGGT 2100
Qy 647 AspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeu 666
Db 2101 GATGTCAAGGGAATCTCCCAAGAGATGAGTCCGCTGAGGCAATGTCTGTGGTCCCTCTA 2160
Qy 667 TrpMetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGly 686
Db 2161 TGGATCGGGATAAGACCTCAATTGAGGAGCAGTGGGAAGATTCTGTCTCCATGCTTGGC 2220
Qy 687 ThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMet 706
Db 2221 ACCCACCCACCTCGCGCAGCTGGACCTGGGCAGCAGCATCTTGACAGAGCGGGCCATG 2280
Qy 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
Db 2281 AAGACCTGTGTGCNAGCTGAGGCATCCCACTTGCAAGATACAGACCTTGATGTTTGA 2340
Qy 727 AsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsn 746
Db 2341 AATGCACAGATTACCCCTGGTGTGCAGCACCTCTGGAGATCGTCATGCGCCCAACCGTAAC 2400
Qy 747 LeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAlaCys 766
Db 2401 CTAAGATCCCTCAACTTGGGAGGACCCACCTCGAAGGAAGAGATGTAAAGGATGCGGTG 2460
Qy 767 GluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeu 786
Db 2461 GAAGCCTTAAACACCCCAAAATGTTTGTGGAGTCTTTGAGGCTGGATGTCTGTGGATTG 2520
Qy 787 ThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSer 806
Db 2521 ACCCATGCTGTACTCAAGATCTCCCAAAATCTTACGACCTCCCCAGCCTGAAATCT 2580
Qy 807 LeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeu 826
Db 2581 CTGAGCCTGGGAGGAAACAAGGTGACAGACCGAGGAGTAATGCCTCTCAGTGATGCCTTG 2640

Qy 827 ArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThr 846
Db 2641 AGAGTCTCCAGTGGCGCTCGCAGAAAGCTGATACTGGAGGACTGTGGCATCACAGCCACG 2700
Qy 847 GlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeu 866
Db 2701 GGTGGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGCACACACCTGTGCTA 2760
Qy 867 SerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuPro 886
Db 2761 TCCAAACAAGCCTGGGGAACGAGGTGTAAATCTACTGTGTCATCCATGAGCCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGGCTGATGTAATCAGTCCACCTGGACACGGCTGGCTGTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGCACCTTGGCTTATGGGTAACTCATGGCTGACGACCTGAGCCTTAGCATGAAC 2940
Qy 927 ProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
Db 2941 CCTGTGAAGACAAATGGCGTGAAGCTTCTGTGCGAGTCTATGAGAACCATCTTGTTCAT 3000
Qy 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966
Db 3001 CTCAGGACCTGGAGTTGGTAAAGTGTCACTCACCGCCGGTCTCTGAGAGTCTGTCC 3060
Qy 967 CysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
Db 3061 TGTGTGATCTCGAGGAGCAGACACCTGAAGAGCTTGGATCTCACGGAATACTTCCCTGGT 3120
Qy 987 AspGlyGlyValAlaAlaLeuCysGluGlyLysGlnLysAsnSerValLeuThrArg 1006
Db 3121 GAGCGTGGGGTGTCTGCACTGTGCGGGACTGAAGCAAAAGAACAGTGTCTGACGAGA 3180
Qy 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
Db 3181 CTCGGTGAAGCATGTGCACTGACTTCTGATTGTCTGTCAGGCACTCTCTTGGCCCTT 3240
Qy 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGly 1046
Db 3241 TCCTGCAACCGGCATCTGACAGTCTAAAACCTGGTGAGAAATACTTTCAGTCCCAAGGA 3300
Qy 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeu 1066
Db 3301 ATGATGAAGCTGTGTGCGCCTTGTGCTGTCCACGCTTAACCTTACAGATAATTTGGGCTG 3360
Qy 1067 TrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLys 1086
Db 3361 TGGAAATGGCAGTACCTGTGCAAATAAGGAAGCTGCTGGAGGAGTGCAGCTACTCAAG 3420
Qy 1087 ProArgValValIleAspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3421 CCCCGAGTCGTAATTCGCGTAGTTGGCAATCTTTTGTGATGAAGATACCGGTAC 3474

RESULT 3

AX704821
LOCUS
DEFINITION
SEQUENCE 1 from Patent EP1285964.
AX704821
ACCESSION
VERSION
AX704821.1 GI:29561487
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
Weiss, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.
Human mater proteins
Patent: EP 1285964-A 1 26-FEB-2003;
SCHERING AKTIENGESSELLSCHAFT (DE)

AX704821 3926 bp DNA linear PAT 04-APR-2003
Sequence 1 from Patent EP1285964.

AX704821
AX704821.1 GI:29561487
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
Weiss, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.
Human mater proteins
Patent: EP 1285964-A 1 26-FEB-2003;
SCHERING AKTIENGESSELLSCHAFT (DE)

FEATURES Location/Qualifiers
source 1..3926
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 5756.00 Matches: 1102
Percent Similarity: 95.34% Conservative: 2
Best Local Similarity: 95.16% Mismatches: 0
Query Match: 81.37% Indels: 54
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AK704821 (1-3926)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAATCGCTCACCTTTCCAGCTACGGGCTGCAATGGTGCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAAGGAAGAAATTTACAGCAATTTCAAGCAATTTCTAAAGAAAGAAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTTCCACAGTTTGAATTCGAGAAATGCCAACGTGGAAATGTTCTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCCTTGATGATGATTAATGAGAGCATCGCTGGCCCTGGGCTAGCTCCATAGCAATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLys----- 96
Db 241 AACATGAACCTCGAAACCCCTCTCGAGAGAGGACACGGGATGACATGAAAAATTTACACCAAGAA 300
Qy 96 ----- 96
Db 301 GATCCTGAAGCAACGATGATGACCAAGACCAAGCAAGAAAAAGTGCCAGAAAAATAAA 360
Qy 96 ----- 96
Db 361 TATGGCATGACTAAGCTTATCTTGGGGGTGCTGACATCTGACTCGAATAATAAACAC 420
Qy 97 -----LysIleSerGlnAlaMetGluGlnGluGly 106
Db 421 AAGTATGTTGGAATTCATCTCTTTTCAGAAATTTTCAAGACTATGGAACAAAGAGGT 480
Qy 107 AlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAlaMetGluGlnGluGlyAla 126
Db 481 GCCACAGCAGACAGACAGAAACAAGAAATTTCAAGCTATGGAACAAAGAGGTGCC 540
Qy 127 ThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHis 146
Db 541 ACACGACAGACAGACAGAAACACAGGACATGGAGGTGACATGGGACTACAGAGTTCAC 600
Qy 147 ValMetThrLysPheAlaGluGluAspValArgArgSerPheGluAsnThrAlaAla 166
Db 601 GTGATGACCAAAATTCGCTGAGGAGGAGATGTACGTGTAGTTTGAACAACACTGCTGCT 660
Qy 167 AspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArg 186
Db 661 GACTGGCCGGAATTCGAAACAGTTGGCTGGTGTCTTTGATTACAGACCCGGTGGGGCTCCCG 720
Qy 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArg 206
Db 721 CCTCGCAGGTGGTCTTCGCACGGAAGTCAGGAATTTGGGAAATCGGCTCTAGCCAGAGG 780
Qy 207 IleValLeuCysTrpAlaGlnGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 781 ATCGTGCTGTCTGGGGCGAAGGTGGACTCTTACCAGGGAAATGTTCTCTCTACGCTCTCTTC 840

Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 841 CTCCCCGTTTAGAGAGATGACGCGGAAGAGAGAGAGAGAGTGTCTCAGAGTTTTCATCTCCAGG 900
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 901 GAGTGGCCAGACTCCAGAGCTCCGGTACCGAGATCATGTCCCGACAGAAAGGCTGTTG 960
Qy 267 PheIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCys 286
Db 961 TTCATCATTCACGGTTTCGATGACCTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGC 1020
Qy 287 LysAspTrpAlaGluLysGlnProProPheThrIleLeuArgSerLeuLeuArgLysVal 306
Db 1021 AAAGACTGGGCTGAGAAAGCAGCCTCCGTTTACCTCATACGAGTCTGTCTGAGGAAGGTC 1080
Qy 307 LeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLys 326
Db 1081 CTGCTCCCTGAGTCTCTCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAG 1140
Qy 327 SerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIle 346
Db 1141 TCAGAGTCTGTCTCCCCGTTACCTGTAGTTAGAGGAATCTCCGGGGAACAAGAAATC 1200
Qy 347 HisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIle 366
Db 1201 CACTTGTCTCTGAGCGCGGATTTGTGAGCATCAGAAAGACACANGGGTTGCGTGCATC 1260
Qy 367 MetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLysCys 386
Db 1261 ATGAAACAACGCTGAGTGTCTGACAGTGCAGAGTGCCAGGTGCCCGCTGGGCTCTCTCATCTGC 1320
Qy 387 ValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db 1321 GTGGCCCTGCAGCTGCAGGACGTGTGGGGAGAGCGTCGCCCTTCAACAAACGCTC 1380
Qy 407 ThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArg 426
Db 1381 ACAGGCTCGACGCCCTTTTGTGTTTCATCAGCTCACCCCTCGAGCGGTGGTCCGGCGC 1440
Qy 427 CysLeuAsnLeuGluArgValValLeuLysArgPheCysArgMetAlaValGluGly 446
Db 1441 TGTCTCAATCTGGAGGAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGGAGGGA 1500
Qy 447 ValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGlu 466
Db 1501 GTGTGGAATAGGAAGTCAAGTGTGACGGTGCAGACCTCATGGTTCAAGGACTCGGGGAG 1560
Qy 467 SerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGlu 486
Db 1561 TCTGAGCTCCGTGCTCTGTTTTCATGAACATCTTCTCCAGACAGCCACTGTGAGGAG 1620
Qy 487 TyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeu 506
Db 1621 TACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCCTTGTACTAGTGTGA 1680
Qy 507 GluGlyLeuGluLeuGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSer 526
Db 1681 GAGGGCTCGAAATCGAGCCAGCTCTCTGCCCTCTGTAGCTTGAGAAAGCAAGAGGTCC 1740
Qy 527 MetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPhe 546
Db 1741 ATGAGACTTAAACAGGACGAGCTTCCATATCCACTCGCTTGGATGAAGGCTTCTTGTGTT 1800
Qy 547 GlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValPro 566
Db 1801 GGCCTCGTGGCGAAGACCGTAGGAGGCCACTGGAGTCTCTGCTGGGCTGTCCCGTTCCT 1860
Qy 567 LeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAla 586
Db 1861 CTGGGGGTGAAGCAGAAAGCTTCTGCACCTGGGTCTCTCTGTGGGTGAGCAGCGCTAATGCC 1920
Qy 587 ThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGlu 606

Db 1921 ACCACCCAGGAGACACCTCGACCCCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAG 1980
Qy 607 PheValArgLeuAlaLeuAAsnSerPheGlnGluValTrrPLeuProIleAAsnGlnAAsnLeu 626
Db 1981 TTTGTTCCTTGGCAATTAACAGCTTCCAGAGAGTGTGGCTTCGATTAACACAGAACCTG 2040
Qy 627 AspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgVal 646
Db 2041 GACTTGATAGCATCTTCCTCTCGCTCCAGCAGCTGTCCGTATTTGGCGGAAATTCGGGTG 2100
Qy 647 AspValLysGlyIlePheProArgAspGlnSerAlaGluAlaCysProValValProLeu 666
Db 2101 GATGTCAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGCATGTCTGTGTGCTCCCTCTA 2160
Qy 667 TrpMetArgAspLysThrLeuIleGluGlnTrrPLeuAspPheCysSerMetLeuGly 686
Db 2161 TGGATCGGGATAGACCTCTATTGAGGAGCAGTGGGAGATTTCTGTCTCCATGCTTGGC 2220
Qy 687 ThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMet 706
Db 2221 ACCCACCCACACCTGCGGCAGCTGGACCTGGGCAGCAGCATCTGCACAGAGCGGCCCATG 2280
Qy 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
Db 2281 AAGACCTCTGTGCGCAAGCTGAGGCATCCACCTGCAAGATACAGACCTCTGATGTTAGA 2340
Qy 727 AsnAlaGlnIleThrProGlyValGlnHisLeuTrrPArgIleValMetAlaAsnArgAsn 746
Db 2341 AATGCACAGATTACCCCTGGTGTGAGCAGCTCTGGAGATCGTCATGCCAACCGTAAC 2400
Qy 747 LeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAlaCys 766
Db 2401 CTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGGAAGAGGATGTAAGGATGGCGTGT 2460
Qy 767 GluAlaLeuLysHisProLysCysLeuLeuGlnSerLeuArgLeuAspCysCysGlyLeu 786
Db 2461 GAAGCCTTAAACACCCAAAATGTTTGTGGAGTCTTTGAGGCTGGATTGTGTGGATTG 2520
Qy 787 ThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSer 806
Db 2521 ACCCATGCTGTACTCGAAGATCTCCCAATPCTTACGACCTCCCGCAGCTGGAATCT 2580
Qy 807 LeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeu 826
Db 2581 CTGACCTGGCAGGAAACAAGGTGACACACGAGGAGTAATGCTCTCAGTGATGCCTTG 2640
Qy 827 ArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThr 846
Db 2641 AGAGTCTCCAGTGGCCCTCGAGAAGCTGATCTGGAGGAGCTGTGGCATCACAGCCACG 2700
Qy 847 GlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeu 866
Db 2701 GTTGTCCAGAGTCTGGCTCAGCCCTCGTACAGACCGAGGCTTGACACACTGTGCGCTA 2760
Qy 867 SerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuPro 886
Db 2761 TCCACACAGCTGGGAGACGAGGTGTAATCTACTGTGTGTCATCATGAGGCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGGCTGATGTGTAATCAGTGCACCTGGACACCGCTGGCTGTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGACATGGCTTATGGGTAACTCATGGCTGACGCACTGAGCCTTAGCATGAAC 2940
Qy 927 ProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
Db 2941 CCTGTGGAGACAAATGGGTGAGGCTTGTGGAGGCTCATGAGAGAACCATCTTGTTCAT 3000
Qy 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966

Db 3001 CTCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCCCGCGTCTGTGAGAGTGTCTGCC 3060
Qy 967 CysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
Db 3061 TGTGTGATCTCGAGAGCAGACACCTCGAAGAGCTTGGATCTCAGGACAATGCCCTGGGT 3120
Qy 987 AspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArg 1006
Db 3121 GACGGTGGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Qy 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
Db 3181 CTCGGGTGAAGGCATGTGACTGACTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGly 1046
Db 3241 TCTGTCAACCGGCATCTGACCACTCTAATCTGGTGACAGATTAATCTCAGTCCCAAGGA 3300
Qy 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeu 1066
Db 3301 ATGATGAAGCTGTGCTGGCTTTCCTGCTGCCACGCTTAATCTACAGATAATTTGGCTG 3360
Qy 1067 TrpLysTrpGlnTrrProValGlnIleArgLysLeuGluGluValGlnLeuLys 1086
Db 3361 TGGAAATGGCAGTACCTGTGCAATTAAGGAAGCTGTGGAGGAAGTGCAGCTACTCAAG 3420
Qy 1087 ProArgValValIleAspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3421 CCCGAGTCTGTAATGACGCTAGTTGGCANTCTTTTGTGATGAGATGACCGGTAC 3474

RESULT 4
AY054986 3885 bp mRNA linear PRI 02-APR-2002
LOCUS Homo sapiens maternal-antigen-that-embryos-require protein (MATER)
DEFINITION mRNA, complete cds.
ACCESSION AY054986
VERSION AY054986.1 GI:19882272
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3885)
AUTHORS Tong Z.B., Bondy C.A., Zhou J. and Nelson L.M.
TITLE A human homologue of mouse Mater, a maternal effect gene essential for early embryonic development
JOURNAL Hum. Reprod. 17 (4), 903-911 (2002)
MEDLINE 21922687
PUBMED 11925379
REFERENCE 2 (bases 1 to 3885)
AUTHORS Tong Z.-B., Bondy, C.A. and Nelson, L.M.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD, NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA

FEATURES
source
1. 3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 3885
/chromosome="19"
/gene="MATER"
1. 3603
/gene="MATER"
/note="oocyte-specific protein; similar to Mus musculus Mater encoded by GenBank Accession Numbers AF074018 and AF143559-AF143573"
/codon_start=1
/product="maternal-antigen-that-embryos-require protein"
/protein_id="AAU15549.1"
/db_xref="GI:19882273"
SSQPCIKMEGDKSLTTFSSYGLWCLVQELDKKEEFQTFKELLKKKSSSESTTCSIQFEIE

NANVECLALLLHEYGYASLAWATSIPIFENNRLTLSEKARDDMKRHSPEBPBRTMTD
QGPSKEVPGISQAOQDSATAAEKEQELSOAMEQSGATAAEETEEOEIQSOMEQEGA
TAATEEQGHGDDTDYKSHVMTFAEEEDVRRSFENTAADNPQMTLAGAFSDRWG
FRPRIVLHGKSGIGKALARRIVLCWAQGLYQGMFSYVFFLPVREMQRKSSVTE
FISREWPDSQAPVTEIMSRPERLLFIIDGDDGLSVLNNDTKLCKDWAERQPPFTLIR
SLLRKVLIPESFLITVTVRDTGTEKLKEVWSPRYLLVRGISGEORIHILLERGITGEHQ
KTQGLRAIINNRELLDOCOVPAVGSLLICVALQLODVGVESVAPFNQTLTGLHAAFAFH
QLTPRGVVRRCLENERVWLRFCRMAVEGVWNEKSVPGDDDLVMQGLSESELALFH
MNILLPUSHCEEYTFPHLSLQDFCAALYVILEGLEIEBPCLPYIVBKTUKRSMELKQA
GFHLSLWMLKFFLEGLVEDVRPEVLGCGPLGVKQKLLHWSVLLGQOPNATTPG
DTLDAFLCFETODEKFRVLRALNSFQEWLPFINQNLDLIASFCLQPCPYLRKIRVDV
KGIFFRDEABEACPVFLMRNDKTLIEQWEDFCSMGLGTHPLRQLDLGSSILITERAM
KTLCAKLHPKCTKIOTLMFRNAQITPGVOHLRWIVMANRNLSINLGLSGTHLKEEDVRM
ACEALHPKCLLESILDCDGLTHACYLKIQUILLTSPSLKSLSLAGNKVTDQGVPL
SDALRVQCALKLILDECGITATGCGSLASALVNSRSLTHLCLSNMNPVEGVNLLCE
RSMELPHCSQRLMLNQCHLDTAGCSLALAMGNSWLTHLSLSMNPVEDGVNLLCE
VMRPSCHLODLVLEKHLTAACCESLSVISRHLKSLDLTDNALDGGPVAALCBG
LKQNSVLTIRLGLXACGLTSCCEALSLALSCNRHLTSLNLVQNNFSPKGMWKLCSAF
ACPTSNLQIIGLWKWQYVPQIRKLEEVQLLKPRVVIDSGSWHSFDEDDRYYWKN"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3885
Score: 5741.50 Matches: 1098
Percent Similarity: 96.16% Conservative: 3
Best Local Similarity: 95.90% Mismatches: 3
Query Match: 81.16% Indels: 41
DB: 9 Gaps: 1

US-10-066-521-6 (1-1344) x AY054986 (1-3885)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTTPCysLeuTyrGlu 20
Db 154 ATGGAAGAGACAAATCGCTCACCTTTTCCAGCTACCGGGCTGCAATGGTGTCTATGAG 213
Qy 21 LeuAspLysGluGluPheGlnThrPhelysGluLeuLeuLysLysSerSerGluSer 40
Db 214 CTAGACAAGGAAGAAATTCAGACATTCAGGAATTAATAAGAGAANAATCTTCAGATCG 273
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaLeuValGluCysLeuAlaLeu 60
Db 274 ACCACATGCTCTATTTCACACAGTTTGAATTCAGAAATCCAGAAATGCTCGGCACCTC 333
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrPalaThrSerIleSerIlePheGlu 80
Db 334 CTCCTTGATGATGATATATGAGGACATCGCTGGCCCTGGGCTACGTCATTAGCATCTTTGAA 393
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetMetLysLys----- 97
Db 394 AACATGAACCTTCGNAACCTCTCGAGNAGGACGCGGATGCATGAAAGACATTCACCA 453
Qy 97 ----- 97
Db 454 GAAGATCTCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGAAAAAGTGCACGAAT 513
Qy 98 -----IleSer 99
Db 514 TCACAAGCTGTGCAACAGATAGTGCCACAGCTGCAGAGACAAAGAACAGGAAATTTCA 573
Qy 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnIleSerGln 119
Db 574 CAAGCTATGGAAACAAGAAAGGTGCCACAGCAGCAGCAGACAGACAGAAATTTTCAAA 633
Qy 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnIleGlyGlyAsp 139
Db 634 GCTATGGAACAAGAAAGGTGCCACAGCAGCAGCAGACAGAAACAAGACATGAGGTGAC 693
Qy 140 ThrThrPheTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
Db 694 ACATGGGACACTAAGAGTACGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCGT 753
Qy 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db 1600 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179

Db 754 AGTTTTGAAAACACACTGCTGCTGACTGGCGGAAAAATGCAAAACGTTGGCTGGTCTTTTGAT 813
Qy 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGGCTTCCGGCCCTCGCACGCTGGTCTTGACGCGAAAAGTCAGAAATTTGG 873
Qy 200 LysSerAlaLeuAlaArgArgIleValLeuLeuCysTrpAlaGlnGlyLeuTyrGlnGly 219
Db 874 AAATCGGCTCTAGCCAGAAAGATCGTGTCTGTGCGGCGCAAGGTGACTCTACACAGGA 933
Qy 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTACCTTCTTCTCTCCCGTTAGAGAGATGCAGCGGAAGAGAGAGCAGT 993
Qy 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTCACAGATTCTCCAGGAGTGCACAGACTCCCGGCTCCGGTGCAGGAGATCATG 1053
Qy 260 SerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspAspLeuGlySerValLeu 279
Db 1054 TCCCGACCAAGAAAGGCTGTGTTCATCATTCAGCGTTTCGATACCTGGGCTCTGTCTC 1113
Qy 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
Db 1114 AACATGACACAAAGCTCTGCAAGACTGGGCTGAGAAGCAGGCTCGGTTCACTCAT 1173
Qy 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 CGCAGTCTGCTGAGGAAGGTCTGCTCCCTGAGTCTCTCTGATCGTCAACCGTCAGAG 1233
Qy 320 ValGlyThrGluLysLysSerGluValValLysProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGCACAGAGAAGCTCAAGTCAGAGGTCTGTCTCCCGGTTACCTGTGTAGTAGGA 1293
Qy 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
Db 1294 ATCTCCGGGGAACAAGAAATCCACTTGTCTCTTGAGCGGGATTGGTGAGCATCAGA 1353
Qy 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
Db 1354 ACAAAAGGTGGTGGCATCATCAACACCGTGAGCTGCTCAGCAGGTGCCAGGTGCCAG 1413
Qy 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerVal 399
Db 1414 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGAGAGCTGGTGGGGAGAGCGCT 1473
Qy 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
Db 1474 GCCCTTCAACCAACGCTCACAGGCTGCAGCGCTTTTTCGGTTTCATCAGCTCACC 1533
Qy 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
Db 1534 CCTCGAGGCGTGGTCCGGGCTGTCTCAATCTCGAGGAAGAGTTGTCTCTGAAGCGCTTC 1593
Qy 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
Db 1594 TGGCGTATGCTGTGAGGAGGTGTGAAATAGGAAGTCAAGTGTGTATGTGTGACGACCT 1653
Qy 460 MetValGlnGlyLeuGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1654 ATGTTCAAGGACTCGGGGAGTCTGAGCTCGGTGCTCTGTTCACATGAACATCTCTC 1713
Qy 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
Db 1714 CCAGACGACGACTGTGAGGAGTACTACCTCTCCACCTCTCCAGCTCTCCAGGACTCTGT 1773
Qy 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
Db 1774 GCCGCTTGTATGATGCTGTGAGGCGCTGGAATCGAGCAGCTCTCTGCCCTCTGTAC 1833
Qy 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
Db 1834 GTTGAGAAGACAAAGAGGTCCATGAGGCTTAAACAGGCGAGCTTCCATATCCACTCGCT 1893

```
Qy 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTTCTTGTGGGCTCGTGAGCGAAGACGTAAGGAGGCCACTGGAGTC 1953
Qy 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGGTCCTCTG 2013
Qy 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGGTGAGCAGCTTATGCGCACACCCAGAGACACCCCTGGAGCGCTTCCACTGTCCT 2073
Qy 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
Db 2074 TTCGAGACTCAAGACAAAGAGTTTGTTCGCTGGCATTTAAACAGCTTCCCAAGAGTGTGG 2133
Qy 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCCGATTAACCAAGAACCTGGACTTGTATGACATCTTCTTCTGCTCCAGCACTGTCCG 2193
Qy 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
Db 2194 TATTTCGGGAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCCGCTGAG 2253
Qy 660 AlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGlu 679
Db 2254 GCATGTCTGTGTCCCTCTATGATGCGGATTAAGACCCCTCATTTGAGGAGCAGTGGAA 2313
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCGTCTCATGTGGCACCCACCCACACCTGGCGGACGTGGACCTGGGGCAGCAGC 2373
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCTTGACAGACGGGCGCATGAGACCTGTGTGCCAAGCTGAGGCATCCCACTGCAAG 2433
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGTATTTAGAAATGCACAGATTACCTCCCTGCTGTGCAACACCTCTCGGA 2493
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGlu 759
Db 2494 ATCTGTCAATGGCCAAACCGTTAATCAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGAA 2553
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGTAGGATGGGTGTGAAGCCTTAAACACCCAAATGTTTGTGGAGTCTTTG 2613
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGGATTGCTGGATTGACCATCGCTGTACCTGAAGATCTCCCAATCCTTTACG 2673
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 2674 ACCTCCCGCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAGGTGACAGCACCCAGGAGTA 2733
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 2734 ACGCTCTCAGTATGCTTGAGGGTCTCCAGTGGCGCCCTGCGAAGAGCTGATCTGGAG 2793
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2794 GACTGTGGCATCACAGCACCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 2853
Qy 860 SerLeuThrHisLeuLysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2854 AGCTTGACACACCTGTGCTATCAACAACAGCCTGGGGAACGAAGGTGTAATCTTACTG 2913
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2914 TGTGATCCATGAGGCTTCCCTCACTGTAGTCTGAGAGGCTGATGCTGAATCAGTGGCCAC 2973
```

```
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2974 CTGACACGCGCTGCTGCTGCTTCTTTCGCACTTATGGCGCTTATGGGTAACCTCATGCTGACG 3033
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 3034 CACTGAGCCTTAGCATGAACCCCTGTGGAAGACAATGGCGTGAAGCTTCTGTGCGAGGTC 3093
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 3094 ATGAGAGAACATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAGTGTCATCTCACCGCC 3153
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 3154 GCGTGTCTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 3213
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 3214 CTCACGACAATGCCCTGGGTGACGGTGGGGTGTCTGCGCTGTGCGAGGAGCTGAAGCAA 3273
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 3274 AAGACAGTGTCTGACGAGACTCGGGTTGAAGGCACTGAGACTGCTCTGATTGCTGT 3333
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 3334 GAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCATCTAAACCTGGTGCAG 3393
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 3394 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGGCCTTGTGCTGTCCACGCTCT 3453
Qy 1060 AsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 3454 AACTTACAGATAATGGGCTGTGGAATGGCAGTACCCCTGTGCAATAAGAAAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAAGTGCGAGCTACTCAAGCCCGAGTGTGTAATTGACGGTAGTTGGCATTCTTTTGTAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588
RESULT 5
LOCUS AX427610 Homo sapiens (human) 3900 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 23 from Patent WO0232955.
ACCESSION AX427610
VERSION AX427610.1 GI:21537730
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Nelson, L.M. and Tong, Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 23 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source
1. .3900
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .3603
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD35973.1"
/db_xref="GI:21537731"
/db_xref="RENTREMBL:CAD35973"
/translation="MKVAGGLELGAALIASPRALVTLSTGPTCSILPKNLPFQNL
SSQPCIKMGDKSLTFSSYGLQWCLYELDKKEEFQTKELLKKKSSSSTTCSIQPFIE
```

NANVECLALLHYYGASLAWATSISIFENNNLRLTSEKARDMDKRHSPEDPATWTD
QGPSKEVPGISQAVQDSATAAEKQEISQAMEQEGATAAEFEQESIQAMEQEGA
TAATEQSGGDTWDTKSHVMTKFAEEDVRRSFENTAADNPEMQTLAFAFSDRWG
FRPRTVHQSGGIGKSGALARRIVLCWAQGLYQGMFSYFFFLFVREMQRKSSVTE
FISREWPDQAPVTEINSRPERLLFIIDGDDLSVLNNDTKLCKOWAEKQPPFTLIR
SLARKVLPSFLITVTVRVGTKEKLVSEVSPRYLLVRGISGEORHLLLERGIGBHO
KTQGLRALINRELLDOCOVPAVGSLLICVALQLODVVGESVAPENQTLTGLHAAPAP
QLTPRGVVRCLNLEVRVLRKFRCMAVEGVWRKSVFDDGDDLMVQGLGSELRALPH
MNILLPDSHCEEYTFPHLSQDFCAALYYVLEGLEIEPALCPLYVEKTRKSMELKQA
GPHHSJLWKRFLFLGLYSEVDRRPLEVLLGCPVLGVQKLLHWSLLGQOPNATPG
DTLDAFCLFETODKEFVRALANSFQEWLPIINONDLIIASSFCLOHCPYLRKIRIVP
KGIIPRDESAEACPVVLMWRDKTLIEQWEDFCSLMGTGTHPHRLQDLGLSSILTERAM
KTLCAKLHPCTQTLQWLFNAQITPGVQHLWRIWMAERNLRSINLGGTHLKEEDVRM
ACEKALPKLLESRLDCCLTHACYIKISQILITSPSLKSLAGNKVTDQGVTP
SDALVSOCAQLQILDECGITATGCOSLASALVNNRSLTHLCLSNNSLNEGWNLC
RSMRLPRCSQRLMNQCHLDTAGCGSLALAGNSWLTLSLSMPNVEDNGVKLICE
VMREPSCHLODLVLCGLTAAACCESLSVCISVRHLKLSLDLTDNALGDGVAAALCEG
LKQNSVLTIRGLKACGLTSDCCALSLSLNRHLTSLNLVQNNFSPKGMKMLCSAF
ACPTSNLIQIIGLWQVYVQIRKLLLEBVLQKLRVVIDGSHWSFDEDDRTWKN"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3900
Score: 5741.50 Matches: 1098
Percent Similarity: 96.16% Conservative: 3
Best Local Similarity: 95.90% Mismatches: 3
Query Match: 81.16% Indels: 41
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AX427610 (1-3900)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTTPCysLeuTyrGlu 20
Db 154 ATGAAGGAGACAAATCGCTCACCTTTTCAGACTACGGCTGCAATGGTGTCTATGAG 213
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 214 CTAGACAGGAGAAATTTACAGCATTTCAAGGAATTTACTAAGAAGAAATCTTCAGAAATCG 273
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 274 ACCACATGCTCTATTCCACAGTTTGAATCGAGAAATCGAGAAATGCTTGGCACTC 333
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTTPAlaThrSerIleSerIlePheGlu 80
Db 334 CTCTTGATGAGTATTATGAGACATCGCTGGCTGCGCTACGTCATTAGCATCTTTGAA 393
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys----- 97
Db 394 AACATGAACCTGCGAACCCCTCTCGAGAGAGGCACGGATGACATGAAGACATTACCA 453
Qy 97 ----- 97
Db 454 GAAGATCTCGAAGCAACGATGACTGACCAAGACCACGAAGAAAAAGTGCCAGGAATT 513
Qy 98 -----IleSer 99
Db 514 TCACAAGCTGTCACAAAGATAGTGCACAGCTGCAGAGACAAAGAACAGGAAATTTCA 573
Qy 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGluIleSerGln 119
Db 574 CAAGCTATGGNACAAAGAGGTGTCACAGCAGCAGCAGCAGACAGACAAAGAAATTTCAAA 633
Qy 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAsp 139
Db 634 GCTATGGAACAAGAGGTGTCACAGCAGCAGCAGCAGACAGAAACAAGGACATGGAGGTGAC 693
Qy 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
Db 694 ACATGGGACATACAAAGTTCACGTGATGACCAATTCGCTGAGGAGGAGGATGACGTCGT 753
Qy 160 SerPheGluAsnThrAlaAlaAspTyrProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db 160 SerPheGluAsnThrAlaAlaAspTyrProGluMetGlnThrLeuAlaGlyAlaPheAsp 179

Db 754 AGTTTTGAAACACTGCTGCTGACTGGCCGGAATGCAAAACGTGTGGCTGGTCTTTGAT 813
Qy 180 SerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGGCTTCGGCCCTCGCACGGTGGTCTGCGCGGAAAGTCAGGAATTGGG 873
Qy 200 LysSerAlaLeuAlaArgArgIleValLeuCysTyrPheAlaGlnGlyLeuTyrGlnGly 219
Db 874 AAATCGGCTTACCCAGAAAGGATCGTGTGTGTCGGCGCAAGTCGACTCTTACCAAGGA 933
Qy 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTACGTCCTTCCTCCCTCCCGTTAGAGAGATGCGAGCGGAAGAAGAGAGCAGT 993
Qy 240 ValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTCACAGAGTTTCATCTCCAGGAGTGCACAGACTCCAGGCTCCGCTGACGAGATCATG 1053
Qy 260 SerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspAspLeuGlySerValLeu 279
Db 1054 TCCCGACCAAGAAAGGCTGTTGTTTCATCATGACGGTTCGATGACCTGGGCTCTGCTCTC 1113
Qy 280 AsnAsnAspThrLysLeuCysLysAspTyrAlaGluLysGlnProProPheThrLeuLeu 299
Db 1114 AACATGACACAAAGCTCTGCAAGAGTGGGCTGAGAAGCAGGCTCGGTTCAACCTCATTA 1173
Qy 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 CGCAGCTGCTGAGGAAGGTCCTGCTCCCTGAGTCTCTGATCGCTCACCGTCAGAGAC 1233
Qy 320 ValGlyThrGluLysLysSerSerGluValValValSerProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGACACAGAGAAAGCTCAAGTCAGAGGTCGTGTCTCCCGGTTACCTGTGTAGTAGGA 1293
Qy 340 IleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
Db 1294 ATCTCCGGGGAACAAGAAATCCACTTGTCTTGTAGCGGGGATTTGGTAGCATCAGAAG 1353
Qy 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
Db 1354 ACACAAAGGTTGCTGCGATCATCAACACCGTGAGCTGTCTCGACCAAGTGCACAGGTGCC 1413
Qy 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerVal 399
Db 1414 GCGGTGGGCTCTCTCATCTGCGTGGCCCTGCAAGTCGAGACGCTGGTGGGGAGAGCGCT 1473
Qy 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
Db 1474 GCGCCCTTCAACCAAGCTCACAGGCTGCAGCGCTGTCGCGTCTTTCGCGTTTCATCAGCTCACC 1533
Qy 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
Db 1534 CCTCGAGGGGTGTCGGGGCTGTCTCAATCTGGAGGAAGAGTTGTCTCTGAAAGCGCTTC 1593
Qy 440 CysArgMetAlaValGluGlyValTyrAsnArgLysSerValPheAspGlyAspAspLeu 459
Db 1594 TGCCTGATGCTGTGAGGAGGTGTGGAATAGGAAGTCAGTGTGTGATGTGACGACCTC 1653
Qy 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1654 ATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTTCACATGAACATCTTCTC 1713
Qy 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
Db 1714 CCAGACAGCCACTGTGAGGAGTACTACACTTCTCCACTCTCCAGTCTCCAGGACTTCTGT 1773
Qy 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
Db 1774 GCGGCTTGTACTACGTGTGTAGGGGCTGGAATCGAGCAGCTCTCTGCTCTCTGTAC 1833
Qy 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
Db 1834 GTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGCAGGCTTCCATATCCACTCGCTT 1893

Qy 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTTCTGTTGGCCCTGCGGAGGAGCGTAAGAGGCGCACTGAGGTC 1953
Qy 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTCGCCCTGCGGGTGAAGCAGAAAGCTTCTGCACTGGGCTCTCTG 2013
Qy 580 LeuGlyGlnGlnProAsnAlaThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGGTGAGCAGCTTAATGCCCACACCCAGAGACACCTTGGACGCCCTTCCACTGCTT 2073
Qy 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
Db 2074 TTCGAGACTCAAGACAAAGAGTTTGTTCGCTGGCATTAACAGCTTCCAAAGAGTGTGG 2133
Qy 620 LeuProLleAsnGlnAsnLeuAspLleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCCGATTAAACAGACCTGGACTTGTATAGCATCTTCTCTGCTCCAGCACTGTCCG 2193
Qy 640 TyrLeuArgLysIleArgValAspValLysGlyLlePheProArgAspGluSerAlaGlu 659
Db 2194 TATTTGCGGAAATTCGGGTGATGTCAAAGGGATCTTCCCAAGAGATGATCCGCTGAG 2253
Qy 660 AlaCysProValProLeuTrpMetArgAspLysThrLeuLleGluGluGlnTrpGlu 679
Db 2254 GCATGTCCTGTGGTCCCTATGATGCGGATAGACCCCTCATTTAGGAGCAGTGGGA 2313
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCTGCTCATGCTTGGCACCCACCCACACCTTGGCGGAGCTGGACCTGGCGAGCAGC 2373
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCCCTGACAGACGGCGGCATGAAGACCTGTGTGCCAAGCTGAGGACATCCCACTGCAAG 2433
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGATGTTTAGAAATGCACAGATTACCCCTGCTGTGCAACACCTCTGGAGA 2493
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 2494 ATCGTCATGGGCAACCGCTAAGATCCCTCAACTTGGGAGGCCACCCACCTGAAGGAA 2553
Qy 760 GluAspValArgMetAlaCysGluAlaLleLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGAAGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTGTGGAGTCTTTG 2613
Qy 780 ArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGGATTGCTGGATTGACCATGCTGTACCTGAAGATCTCCCAATCTTACG 2673
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 2674 ACCTCCCCAGCCTGAATCTCTGAGCTGGCAGGAAACAAGGTGACAGACAGGAGTA 2733
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuLleLeuGlu 839
Db 2734 AGCCCTCTCAGTGATGCTTGAAGGTCTCCAGTGGCGCTTGCAGAAAGCTGATACTGGAG 2793
Qy 840 AspCysGlyLleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2794 GACTGTGGCATCAGACAGCAGCGGTTGCCAGAGTCTGGCTTACGCCCTCTGACGACCCG 2853
Qy 860 SerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2854 AGCTTGACACACTGTGCTATCCCAACAACAGCTTGGGGAACGAAGGTGTAATCTACTG 2913
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2914 TGTGATCCATGAGGCTTCCCCACTGTATGTGACAGAGGCTGATGCTGAATCAGTGGCCAC 2973

Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2974 CTGGACACGGCTGGCTGTGGTTCTCTTGCACTTGGCTTATGGTAACCTATGGCTCAGC 3033
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuGluVal 939
Db 3034 CACTGAGCCTTAGCATGAACCCCTGTGGAAGACAATGGCTGAAGCTTCTGTGGAGGTC 3093
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 3094 ATGAGAACCACTTGTGTCATCTCCAGACCTGGAGTTGGTAAAGTCTCATCTCACGCC 3153
Qy 960 AlaCysCysGluSerLeuSerCysValLleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 3154 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 3213
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 3214 CTCAGGACAATGCCCTGGGTGACGGTGGGTGTGCTGCTGTCGAGGAGCTGAAGCAA 3273
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCys 1019
Db 3274 AAGAACAGTGTCTGACGACACTCGGTTGAGGCGATGTGGACTGACTTCTGATTGCTGT 3333
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 3334 GAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTGTAAACCTGGTGCAG 3393
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 3394 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGCGCCTTGTGCTGCCACGCTCT 3453
Qy 1060 AsnLeuGlnIleLleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 3454 AACTTACAGATTAATGGGCTGTGGAATGCGATGACCTGTGCAAAATGAGGAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLysProArgValValLleAspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAGTGCAGCTACTCAAGCCCGAGTCTTAATTCAGCGTAGTGTGGCTTCTTTTGTAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588
RESULT 6
AX459881 AX459881 5859 bp DNA linear PAT 30-NOV-2002
LOCUS Sequence 24 from Patent WO0240668.
DEFINITION AX459881
ACCESSION AX459881
VERSION AX459881.1 GI:21725649
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Techopp, J. and Martinon, P.
TITLE Proteins and dna sequences underlying these proteins used for
treating inflammations
JOURNAL Patent: WO 0240668-A 24 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source
1. 5859
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP8/Py12---Py8.dna"
ORIGIN
Alignment Scores: 0 Length: 5859
Pred. No.: 5724.50 Matches: 1106
Score: 90.53%
Percent Similarity: 90.53%

His 15 m
Camm

Best Local Similarity:	90.29%	Mismatches:	7
Query Match:	80.92%	Indels:	109
DB:	6	Gaps:	3
US-10-066-521-6 (1-1344) x AX459881 (1-5859)			
QY	1	MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu	20
DB	1975	ATGGAAGAGACAAATCGCTCACCTTTTCCAGCTACCGGCTGCAATGGTGTCTATGAG	2034
QY	21	LeuAspLysGluGluPheGlnThrPhelysGluLeuLeuLysLysSerSerGluSer	40
DB	2035	CTAGACAAGGAAGAAATTCACACATTCAGGAATTAATAAGAGAANAATCTTCAGAAATCG	2094
QY	41	ThrThrCysSerIleProGlnPheGluIleGluAlaSerLeuAlaValGluCysLeuAlaLeu	60
DB	2095	ACCACATGCTCTATTTCACAGTTTGAATTCAGAAATCCAGAAATCCAAAGTGAATGTCTGGCAGTC	2154
QY	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu	80
DB	2155	CTCTTGATGAGTATATATGGAGCATCGCTGGCCCTGGGCTACGTCCATTAGCATCTTTGAA	2214
QY	81	AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys-----	97
DB	2215	AACATGAACCTGCGAACCTCTCGAGAGAGGACGCGATGACATGAAAGACATTCACCA	2274
QY	97	-----	97
DB	2275	GAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCACGAANAAT	2334
QY	98	-----	98
DB	2335	TCACAAGCTGTGCACAAGATAGTGCACAGCTCGAGAGACAAAGAACCAAGAAATTTCA	2394
QY	100	GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln	119
DB	2395	CAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAGAGAACAAGAAATTTCCAA	2454
QY	120	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyAsp	139
DB	2455	GCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGAGAACAAGACATGGAGGTGCAC	2514
QY	140	ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg	159
DB	2515	ACATGGGACTACAAGTTCAGTATGACCAATTCGCTGAGGAGGAGATGTACGTCGT	2574
QY	160	SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp	179
DB	2575	AGTTTTGAACAACATGCTGCTGACTGCGCGGAATGCAACGTTGGCTGGTCTTTTGTAT	2634
QY	180	SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly	199
DB	2635	TCAGACCGGTGGGCTTCCGCGCTCGCACGCTGGTCTTGACGCGAAAGTCAGGAATTTGGG	2694
QY	200	LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly	219
DB	2695	AAATCGGCTCTAGCCAGAAAGATCGTCTGTGCTGGGCGCAAGGTGACATCTACAGGGA	2754
QY	220	MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer	239
DB	2755	ATGTTCTCTACGCTCTTCTCTCTCCCGTTAGAGAGATGACGCGGAAGAGGAGAGCAGT	2814
QY	240	ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet	259
DB	2815	GTCACAGATTTCACTCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGCTGACCGAGATCATG	2874
QY	260	SerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeu	279
DB	2875	TCCCGACAGAAAGCTGTGTTTCAATGACCGGTTTCAGTACCTGGGCTCTGTCTCTC	2934
QY	280	AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle	299
DB	2935	AACAAATGACACAAAGCTCTGCAAGAGACTGGGCTGCAAGAGCAGCCTCGCTTCAACCTCATA	2994

QY	300	ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuLeuValThrValArgAsp	319
DB	2995	CGCAGTCTGCTGAGGAAGGTCTCTGCTCCCTGAGTCTCTGATCGTCAACCGTCAGAGAC	3054
QY	320	ValGlyThrGluLysLeuLysSerGluValValLysProArgTyrLeuLeuValArgGly	339
DB	3055	GTGGGCACAGAGAAGCTCAAGTCAGAGGTCTGTCTCCCCGTTACCTGTTAGTTAGAGGA	3114
QY	340	IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys	359
DB	3115	ATCTCCGGGGAACAAAGAATCCACTTGTCTCTTGAGCGGGATTTGCTGAGCATCAGAAG	3174
QY	360	ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro	379
DB	3175	ACAAAGGGTTGCTGCGATCATGAACAAACCGTGAGTGTCTCGACCAAGTGCAGGTGCC	3234
QY	380	AlaValGlySerLeuLysCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal	399
DB	3235	GCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGAGGACGTGGTGGGGAGAGCGTC	3294
QY	400	AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr	419
DB	3295	GCCCCCTTCAACCAACAGCTCACAGGCTGCACGCGCTTTTGTGTTTCATCAGTCACCC	3354
QY	420	ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe	439
DB	3355	CCTCGAGGGCTGTCTCGGCGCTGTCTCAATCTGGAGGAAGAGTTGCTCTGAAAGCGCTTC	3414
QY	440	CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu	459
DB	3415	TGCGGTATGCTGTGGAGGAGTGTGGAAATAGGAAGTCAAGTGTTCACGCTGACGACCTC	3474
QY	460	MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu	479
DB	3475	ATGTTTCAAGGACTCGGGGAGTCTGAGCTCGGTCTCTGTTTTCACATGAACATCTTCTC	3534
QY	480	ProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAspPheCys	499
DB	3535	CCAGACGCGACTGTGAGGAGTACTACAGCTTCTCCACCTCAGCTCTCCAGGACTTCTGT	3594
QY	500	AlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr	519
DB	3595	GCGGCTTGTACTACGTGTAGAGGCGCTGGAATTCAGCCAGCTCTCTGCGCTCTGTAC	3654
QY	520	ValGluLysThrLysArgSerMetGluLeuLeuLysGlnAlaGlyPheHisIleHisSerLeu	539
DB	3655	GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGGCTTCCATATCCACTCGCTT	3714
QY	540	TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal	559
DB	3715	TGGATGAAGCGTTTCTGTTTGGCTCTGAGCGAAGACGTAAAGAGGCGCACCTGAGGTC	3774
QY	560	LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu	579
DB	3775	CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGACCTGGGTCTCTCTG	3834
QY	580	LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu	599
DB	3835	TTGGGTCAGCAGCTTAATGCCACCCAGGAGACACCTCGGACGCTTCCACTGCTCTT	3894
QY	600	PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp	619
DB	3895	TTCCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACAGCTTCCAAAGAGTGTGG	3954
QY	620	LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro	639
DB	3955	CTTCCGATTAACACAGAACTGGACTTGATAGCATCTTCTTCTGCGCTCCAGCACTGTCCG	4014
QY	640	TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu	659
DB	4015	TATTTTCGGAATAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGATGATGTCGCTGAG	4074


```
Qy 660 AlaCysProValProLeuTriMetArgAspLysThrLeuIleGluGlnTriPglu 679
Db 4075 GCATGCTCTGGTCCCTCTATGATGGCGGATGAAGACCCCTCATTTGAGGACAGTGGAA 4134
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 4135 GATTTCTCTCATGCTTGGCACCCACCCACCTTGGCGGAGCTGGACCTGGGAGCAGC 4194
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 4195 ATCCTGACAGAGGGGCCATGAGACCCCTGTGTGCCAGCTGAGGCATCCACCTGCAG 4254
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTriPArg 739
Db 4255 ATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGCACCTCTGGAGA 4314
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGlu 759
Db 4315 ATCGTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGACCCACCTGAAGAA 4374
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 4375 GAGGATGTAAAGATGGCGTGTGAAGCCTTAATAACACCCCAAAATGTTTGTGGAGTCTTTG 4434
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 4435 AGGCTGGATTGCTGGGATTGACCATGCTTACCTGAAAGATCTCCCAATCCTTTAAG 4494
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 4495 ACCTCCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACACGAGTA 4554
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 4555 ATGCCTCTCAGTGATGCTTCAGAGTCTCCAGTGGCCCTGCAAGAGCTGATACTGAG 4614
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 4615 GACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGCCCTCAGCCCTCGTCAGCAACCGG 4674
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 4675 AGCTTGACACACCTGTGCTTATCCAAACACACCTGGGGAAACGAAGGTGTAATCTACTG 4734
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 4735 TGTGATCCATGAGGCTTCCCACACTGTAGTCTGCAGAGCTGTAGTGAATCAGTGCAC 4794
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTriPLeuThr 919
Db 4795 CTGGACACGGCTGGCTGTGGTTTCTTTCGACTTGGCTTATGGGTAACTCATGCTGACG 4854
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 4855 CACCTGAGCCTTAGCATGAACCCCTGTGGAAGACAAATGGCGTGAAGCTTCTGTGGGAGTC 4914
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 4915 ATGAGAAGAACCATCTTGTCACTCCAGAGCTGGAGTTGGTAAAGTGTCACTCACCGCC 4974
Qy 960 AlaCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 4975 GCGTCTGTGAGAGTCTCTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 5034
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaLeuCysGluGlyLeuLysGln 999
Db 5035 CTCACGGCAATGCCCTGGGTGACGGTGGGTGTGTCACCTGTGCAGGAGGACTGAAGCAA 5094
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 5095 AAGAACAGTGTCTGACAGACTCGGGTTGAAGGCATGTGACACTCTGATTTGCTGT 5154
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
```

```
Db 5155 GAGGCACCTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCAGTCTAAACCTGGTGCAG 5214
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 5215 AATAACTTTCAGTCCCAAGGAAGATGATGAAGCTGTGTTTGGCCCTTTCCTGTGCCACGTCT 5274
Qy 1060 AsnLeuGlnIleIleGly----- 1065
Db 5275 AACTTACAGATANTGGCAATGACTCTGAAGAAAATGACCTTCTTCGAGAATCTGCTCTA 5334
Qy 1065 ----- 1065
Db 5335 GTAGTTTTTCTTAAAGTCACTGTTTCCAGAACCTATCAATGACATTAAGGGAGAACTTA 5394
Qy 1065 ----- 1065
Db 5395 CTGTACCTCCCAACCCCTATAACAGACCCGGGACAGAGACTCTGAGGAAGAAATTCAT 5454
Qy 1066 -----LeuTriPlysTriPLeuTriProValGlnIleArgLysLeuLeuGlu 1080
Db 5455 GGATGGACTGAAGGCTGTGGAATGGCACTACCTGTGCAATTAAGNAAGCTGCTGAG 5514
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTriPHisSerPheAspGlu 1100
Db 5515 GAAAGTGACGCTACTCAAGCCCGAGTGTAAATGACGGTAGTTGGCACTTCTTTTGATGAA 5574
Qy 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 5575 GATGACCGATTGGATTCTTCAAGCCAGCAAAACAGTCACCTCAGCAAGACAGACTTACAAT 5634
Qy 1112 LeuProGluSerArg 1116
Db 5635 CTATGCGCATCAG 5649
RESULT 7
AX459873 LOCUS AX459873 6939 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 16 from Patent WO0240668.
ACCESSION AX459873
VERSION AX459873.1 GI:21725645
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Teschopp,J. and Martinon,F.
TITLE Proteins and dna sequences underlying these proteins used for treating inflammations
JOURNAL Patent: WO 0240668-A 16 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source 1. 6939
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="NALP5/Py8.dna /GENSCAN_predicted_CDS_1/ 6939_bp"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3
US-10-066-521-6 (1-1344) x AX459873 (1-6939)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTriGlyLeuGlnTriPLeuThrGlu 20
Db 3055 ATGGAAGAGACAAATCGCTCACCTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 3114
```


QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuIleLysSerSerGluSer 40
DB 3115 CTAGCAAGGAGAAATTTTCAGACATTTCAAGAAATTTACTAAAGAAAGAAATCTTCAAGATCG 3174
QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
DB 3175 ACCACATGCTCTATTTCCACAGTTTGAATCGAGATGCCACGTGGAAATGTCCTGGCACTC 3234
QY 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
DB 3235 CTCCTTGATGATGATATTATGAGACATCGCTGGCCCTGGCTACGTCATAGCATCTTTGAA 3294
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
DB 3295 AACATGAACCTTGCAGAACCTCTCGAGAAAGGCACGGGATGACATGAAGAAAGACATTACCA 3354
QY 97 ----- 97
DB 3355 GAAGATCTCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCACAGGAATT 3414
QY 98 -----IleSer 99
DB 3415 TCACAAGCTGTGCAACAAGATAGTGCACAGCTGCAGAGACAAAGAAATTTCA 3474
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln 119
DB 3475 CAAGCTATGGAAACAAGAGGTGCCACAGCAGCAGAGACAGAAACAAGAAATTTCAAA 3534
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGlyLys 139
DB 3535 GCTATGGNACAAAGAGTGCACAGCAGCAGCAGACAGACAGACATGGAGGTGAC 3594
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
DB 3595 ACATGGGACTACAAGAGTCAAGTATGATACCAAAATTCGCTGAGGAGGAGGTACGTCGT 3654
QY 160 SerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
DB 3655 AGTTTTGAANAACATGCTGCTGACTGGCCGGAATGCAAAAGTTGGCTGGCTTTTGAT 3714
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
DB 3715 TCACACGGTGGGGCTTCGGCCCTCGCACGGTGGTTCTGACGGAAGATCAGGNAITGGG 3774
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGly 219
DB 3775 AAATCGGCTCTAGCCAGAAAGGATCGTGTGCTGGCGCGCAAGGTGACTCTACCAAGGA 3834
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
DB 3835 ATGTTCCTCTACGCTTCTTCCTCCCGTTAGAGAGATGCAGCGGAAGAGGAGAGCAGT 3894
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
DB 3895 GTCACAGATTCACTCTCAGGGAGTGGCCAGACTCCACGGCTCCGGTGCAGGAGATCATG 3954
QY 260 SerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeu 279
DB 3955 TCCCGACCAAGAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGTCCTC 4014
QY 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
DB 4015 AACAAATGACAAAGAGCTTCGAAAGACTGGGCTGAGAAAGAGCGCTCCGTTCAACCTCAT 4074
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
DB 4075 CGCAGTCTGCTGAGAAAGTCTGCTCCCTGAGTCTTCTGATCTGCACCGTCAGAGAC 4134
QY 320 ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
DB 4135 GTGGGCAACAGAGAGCTCAAGTCAGAGGTCGTGTCTCCCCGTTACCTGTTAGTAGGA 4194

QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 4195 ATCTCCGGGGAAACAAAGAATCCACTTGCTCTTTAGCGGGGATTTGTTAGCATCAGAAG 4254
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
DB 4255 ACACNAGGTTGGTGGGATCATGAACACCGTAGCTGCTCAGCAGTGCCAGGTGCCCC 4314
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399
DB 4315 GCGTGGGCTCTCTCATCTGCTGGCGCTGCAGCTGCAGGACGTGCTGGGGAGAGCGTC 4374
QY 400 AlaProPheAsnGlnThrIleThrGlyIleHisAlaAlaPheValPheHisGlnLeuThr 419
DB 4375 GCCCCTTCAACCAACAGCTCACAGGCTGCACCGCGCTTTTGTGTTTCATCAGCTCACC 4434
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
DB 4435 CCTCGAGGCTGTGCGGCGCTGTCTCAATCTGGAGGAAAGAGTTTGTCTGAAGCGCTTC 4494
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
DB 4495 TGCCTGATGGCTGTGAGGGAGTGTGGAATAGGAAGTCAGTGTTCAGCGTGACGACCTC 4554
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 4555 ATGCTTCAAGGACTCTGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGAACATCTTCTC 4614
QY 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 4615 CCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACTCAGTCTCCAGGACTCTCTGT 4674
QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
DB 4675 GCGCCTTGTACTACTGTTAGAGGCTTGGAAATCGAGCCAGCTCTCTGCCCTCTGTAC 4734
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 4735 GTTGAGAAACAAAGAGGTCCATGGAGCTTAAACAGCGAGGCTTCCATATCCACTCGCTT 4794
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
DB 4795 TGGATGAAGGTTCTTGTGTCCTGTGAGCGAAGACGTAAAGAGGGCCACTGGAGGTTC 4854
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 4855 CTGCTGGGCTGTCCCGTTCCTCCCTGGGGGTGAAGCAAGCTTCTGCACCTGGGTCTCTCTG 4914
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGGTGACGAGCCTAATGCCACCCAGGAGACACCTCGAGCGCTTCCACTGCTCTT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTTCAGACTCAACACAAAGATTTGTTCCCTTGGCATTAAACAGCTTCCAGAGGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCGATTAACACAGAACCTGGACTTGTATAGCATCTTCTTCTGCTCCAGCACCTGCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTCCGAAAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAGATAGTCCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
DB 5155 GCATGTCTGTGGTCCCTCTATGGATGCGGGATAGACCTCTATTGAGAGAGCAGTGGGA 5214
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 5215 GATTTCTGCTCCATGTTGGCACCCACACCTGCGGCAGCTGGACCTGGGCAGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719

Db 5275 ATCTTGACAGACGGCGCCATGAAGACCTGTGTGTCAGCTGAGGCATCCACCTGCAAG 5334
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 5335 ATACAGACCTGTAGTGTAGAAATGACAGATTACCCCTGGTGTGACGACCTCTGGAGA 5394
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 5395 ATCGTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGAA 5454
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 5455 GAGGATGTAAGGATGGCGTGTGAAGCCCTAAACACCCCAAAATGTTGTTGGAGTCTTTG 5514
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysValLeuLysIleSerGlnIleLeuThr 799
Db 5515 AGGCTGGATTGCTGGGATTGACCATGCTGTTACCTGAAGATCTCCCAATCTTACG 5574
Qy 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 5575 ACCTCCCCAGCCTGAATCTCTGAGCCTGCGAGAAACAAAGGTGACAGACCCAGGAGTA 5634
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleGlu 839
Db 5635 ATGGCTCTCAGTGATGCTTGAAGTCTCCAGTGTCCAGTGTGCGCCCTGCAGAAAGCTGTATCTGGAG 5694
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 5695 GACTGTGCATACACGCCCGGTTGCCAGATCTGGCCCTCAGCCCTCGTCAGCAACCGG 5754
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 5755 AGCTTGACACACCTGTGCTTATCAACAAACAGCCTGGGGAACGAAGGTGTAATCTACTG 5814
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 5815 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATCGAATCAGTGCAC 5874
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 5875 CTGACAGCGCTGGCTGTGGTGTCTTCTGCACTTGGCTTATGGGTAACTCATGCTCAGC 5934
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 5935 CACCTGAGCCTTAGCATGAACCTGTGGAACAATGGCGTGAAGCTTCTGTGCGAGTC 5994
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 5995 ATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACGCC 6054
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 6055 GCGTGTGTGAGTGTCTGTCTGTGTGATCTCGAGGACACACACCTGGAAGCGCTGGAT 6114
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 6115 CTCAGGACAATGCCCTGGGTGACGGTGGGTGTCTGCACTGTGCGAGGACTGAAGCAA 6174
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 6175 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGACCTGACTTCTGATTGCTGT 6234
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 6235 GAGGCACCTCTCTTGGCCCTTCTTCTGCAACCGGCATCTGACCACTGATAAACCTGGTCAG 6294
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 6295 AATAACTTCAGTCCCAAGGAATGATGAAGTGTGTTCGGCTTGTGCTGTCCACGCTCT 6354
Qy 1060 AsnLeuGlnIleLeuGly----- 1065

Db 6355 AACTTACAGATAAATGGCAATGACTCTGAAGAAAAATGACGTTCTTCGGAATCTGCTCTA 6414
Qy 1065 ----- 1065
Db 6415 GTAGTTTGTCTAAAGTCACTGTTTCCAGAACCTATCAATGACATTAAGGGAGAACTTA 6474
Qy 1065 ----- 1065
Db 6475 CTGTACCTCCCAAAACCTATAACAGACCCCGGCACAGAGACTCTGAGGAAGGAATTCAT 6534
Qy 1066 -----LeuTrpLysTrpGlnTyProValGlnIleArgLysLeuLeuGlu 1080
Db 6535 GSATGGACTGAAGGCTGTGGAATGCGAGTACCCCTGTGCAAAATAAGGAAGCTGCTGGAG 6594
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 6595 GAAGTGCAGCTACTCAAGCCCGAGTGTAAATGACGCTAGTTGGCATTTCTTTTGATGAA 6654
Qy 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 6655 GATGACCGATTGGATCTTCAAGCCAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 6714
Qy 1112 LeuProGluSerArg 1116
Db 6715 CTCATGCGATCAGAC 6729
RESULT 8
LOCUS AX459891 6939 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 34 from Patent WO0240668.
ACCESSION AX459891
VERSION AX459891.1 GI:21725654
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Techopp,J. and Martinon,F.
TITLE Proteins and dna sequences underlying these proteins used for
treating inflammations
JOURNAL Patent: WO 0240668-A 34 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source 1..6939
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="NALP13/Py17.cdna"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3
US-10-066-521-6 (1-1344) x AX459891 (1-6939)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyArgLeuGlnTrpCysLeuTrpGlu 20
Db 3055 ATGGAAGGAGACAAATCGCTCACCCTTTCAGCTACGGGCTGCAATGGTCTCTATGAG 3114
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysLysSerSerGluSer 40
Db 3115 CTAGCAAGGAAGAAATTCAGACATTCAGAAATTAAGGAAGAAATCTTCAGAAATCG 3174
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 3175 ACCACATGCTCTATTCCACAGTTTGAATTCGAAGATGCCAACCTGGAATGCTGGCAGCTC 3234

QY 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrSerIleSerIlePheGlu 80
DB 3235 CTCCTGATGATGATTATGGAGCATCGCTGGCTGGCTAGCTCAATTAGCATCTTTGAA 3294
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys----- 97
DB 3295 AACATGAACCTGCGAACCTCTCGGAGAGGACACGGGATGACATGAAAAGACATTCACCA 3354
QY 97 ----- 97
DB 3355 GAAGATCCTGAAGCAACGATGACTGACCAGGACCAAGCAAGAAAAAGTGCCAGGAATT 3414
QY 98 -----IleSer 99
DB 3415 TCACAGCTGTGCAACAAGATAGTGCACAGCTGCAGAGACAAAGAACCAAGAAATTTCA 3474
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln 119
DB 3475 CAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGAGAACAAGAAATTTCAAA 3534
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlyHisGlyAsp 139
DB 3535 GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGAGAACAAGACATGGAAGGTGAC 3594
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg 159
DB 3595 ACATGGGACTACAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGATGTACGTGT 3654
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
DB 3655 AGTTTTGAAAACACTGCTGCTGACTGGCCGGAATGCAAAAGTTGGCTGGTCTTTGAT 3714
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
DB 3715 TCACACCGGTGGGGCTTCCGGCTCGCACGGTGTCTTCGACGGAAGTCAGGAAATGGG 3774
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly 219
DB 3775 AAATCGGCTCTAGCCAGNAGATGTGTGTGTGTGGGCGCAAGGTGACCTTACCAGGGA 3834
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
DB 3835 ATGTTCTCTACCTCTTCTCTCTCCCGTTAGAGAGATGCAGCGGAAGAGGACAGCAGT 3894
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
DB 3895 GTCAAGAGTTTCATCTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTGACGAGATCATG 3954
QY 260 SerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeu 279
DB 3955 TCCCGACCAAGAGCTGTGTTTCATCATTCACCGTTTCGATGACCTGGGCTCTGTCTCTC 4014
QY 280 AsnAsnAspThrLysLysCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
DB 4015 AACAAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAAGCAGCTCCGTTCCACCTCATA 4074
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheIleIleValThrValArgAsp 319
DB 4075 CGCAGTCTGTGAGAAAGGTCTCTCTCTGAGTCCCTTCCTGATCGTCACCGTCAGAGAC 4134
QY 320 ValGlyThrGluLysLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
DB 4135 GTGGGCACAGAGAGCTCAAGTCAAGGTCTGTCTCTCCCGTTACCTGTTAGTTAGAGGA 4194
QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 4195 ATCTCCGGGGAACAAGAAATCCACTTGTCTCTGAGCGCGGATTGGTGAGCATCAGAAG 4254
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
DB 4255 ACACAAGGGTGTGCGATCATGAACAACCGTAGCTGTCTGACCCAGTGCAGGTGCC 4314
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399

DB 4315 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGTGTGGGGAGAGCGTC 4374
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
DB 4375 GCCCCCTTCAACCAACGCTCACAGGCTGCACGCGCTTTGTGTTCATCAGCTCACC 4434
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
DB 4435 CCTCGAGGCGTGTGCGGCGCTGTCTCAATCTCGAGGAAGAGTTGTCTCGAAGCGCTTC 4494
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
DB 4495 TGGCGTATGGCTGTGAGGGAGTGTGGAATAGGAAGTCACTGTGTTCAGGTGAGACCTC 4554
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 4555 ATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTCACATGAACATCTTCTC 4614
QY 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 4615 CCAGACAGCCACTGTGAGGAGTACTACCTTCTCCACTCAGCTCTCCAGGACTTCTGT 4674
QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
DB 4675 CGCGCTTGTACTACGTGTAGAGGCGCTGGAATCGAGCCAGCTCTCTGCCCTCTGTAC 4734
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 4735 GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGCAGGCTTCCATATCCACTCGCTT 4794
QY 540 TrpMetLysArgPheIlePheGlyValValSerGluAspValArgArgProLeuGluVal 559
DB 4795 TGGATGAAGGTTCTTGTGTGGCTCTGTAGCGAAGACGTAAAGGAGGCCACTTGGAGGTC 4854
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 4855 CTGCTGGGCTGTCCCGTTCCTCTGGGGGTGAACAGAGAGCTTCTGCACCTGGCTCTCTG 4914
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGGTCAGCAGCCTAATGCCACACCCAGGAGACACCTCGGACGCTTCCACTGTCTT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTCGAGACTCAAGACAAAGAGTTGTTCGCTTGGCATTAACAGACTTCCAAAGAGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCGGATTAACCAAGAACCTGGACTTGTATAGCATCTTCTCTGCTCCAGCAGCTGTCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTGCGGAAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGAGATGATGTCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrIleIleGluGluGlnTrpGlu 679
DB 5155 GCATGCTGTGTCTCTATGATGCGGGAATAAGACCTCATTTGAGGAGCAGTGGGAA 5214
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 5215 GATTTCTGTCTCATGCTTGGCACCCACCCACACTCGGGCAGCTGGACCTGGGCGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
DB 5275 ATCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAG 5334
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
DB 5335 ATACAGACCTGTATGTTTAGAAATGACAGATTTACCCCTGTGTGACGACCTCTTGAGA 5394
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759

```
Db 5395 ATCGTCATGGCCAAACCGTAAGTCCCTCAACTTTGGGAGGACCCACCTGAAGAA 5454
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 5455 GAGGATGTAAGATGGCGTGTGAGCCTTAAACACACCCAAATGTTGTTGAGTCTTTG 5514
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 5515 AGGCTGGATGCTGTGGATGACCCATGCTTACCTGAAGATCTCCCAAATCCTTACG 5574
Qy 800 ThrSerProSerLeuLysSerLeuLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 5575 ACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAAGGTGACACACAGGAGTA 5634
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 5635 ATGCTCTCAGTGATGCTTGGAGTCTCCAGTGGCCCTGCAGAGCTGATACTGGAG 5694
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 5695 GACTGTGGCATCACAGCCAGCGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 5754
Qy 860 SerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 5755 AGCTTGACACACCTGTGCTATCCAAACACAGCCTGGGGAAACGAAGGTGTAATCTACTG 5814
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 5815 TGTGATCATGAGGCTTCCCACTGTGATCTGACAGAGCTGATGCTGAATCAGTGCAC 5874
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 5875 CTGGACAGCGTGGCTGTGTTCTTCTGACCTTGGCTTATGGGTAACTCATGCTGACG 5934
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 5935 CACCTGAGCCTTAGCATGAACCTGTGGAAGCAATGGCTGAAGCTTCTGTGGAGGTC 5994
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 5995 ATGAGAGAACCACTTGTGTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCACTCACCCG 6054
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 6055 GCGTCTGTGAGAGTCTGCTGCTGTGATCTCGAGAGCAGACACCTGAAGAGCCTGGAT 6114
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 6115 CTCACGGACAATGCCCTGGGTGACGGTGGGGTTGCTGCACCTGTGCGAGGAGCTGAAGCA 6174
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 6175 AGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGT 6234
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 6235 GAGGCACTCTCTTGGCCCTTCTGTCACACCGCATCTGACCACTTAAACCTGGTGAG 6294
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 6295 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGTGGCCCTGTCCCACTGT 6354
Qy 1060 AsnLeuGlnIleIleGly----- 1065
Db 6355 AACTTACAGATAATGGCAATGACTCTGAAGAAAATGACGTTCTTCGAGAAATCTGCTCTA 6414
Qy 1065 ----- 1065
Db 6415 GTAGTTTGTCTAAGTCACTGTTTCCAAGAACCTTATCAATGACATTAAGGAGAACTTA 6474
Qy 1065 ----- 1065
Db 6475 CTGTACTCTCCCAAAACCTTATAACAGACCCGCGCACAGAGACTCTGTGAGGAAGGAATTCAT 6534
```

```
Qy 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 6535 GGATGGACTCAAAGGCTGTGAAATGCGAGTACCCCTGTGCAATTAAGAAAGCTGCTGAG 6594
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 6595 GNAAGTCAGCTACTCAAGCCCGAGTGTAAATGACGGTAGTTGGCAATCTCTTTGATGAA 6654
Qy 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 6655 GATGACCGATTTGATCTTCAAGCCCAAAACAGTCACTCAGCAACAGACTTACAAT 6714
Qy 1112 LeuProGluSerArg 1116
Db 6715 CTCATGGCATCACAG 6729

RESULT 9
AX704823 3830 bp DNA linear PAT 04-APR-2003
LOCUS Sequence 3 from Patent EP1285964.
DEFINITION AX704823
ACCESSION AX704823
VERSION AX704823.1 GI:29561488
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Weiss, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.
TITLE Human mater proteins
JOURNAL Patent: EP 1285964-A 3 26-FEB-2003;
SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES
source 1. .3830
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. NO.: 0 Length: 3830
Score: 5693.50 Matches: 1087
Percent Similarity: 95.87% Conservative: 5
Best Local Similarity: 95.43% Mismatches: 12
Query Match: 80.34% Indels: 35
DB: Gaps: 2

US-10-066-521-6 (1-1344) x AX704823 (1-3830)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGAGACAATCGCTCACCTTTCCAGCTACGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTACACAAGGAAGATTTACAGACATTCAGGAATTTCTAAAGAAGAAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATVTCACACATTTGAAATCGAGAAATCGCAACGTCGGAATGTCTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCCTTGATGAGTATATGAGACATCGCTGGCTTGGCTACCTCCATAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysIleSerGln 100
Db 241 AACATGAACCTCGCAACCTCTCGAGAGAGCGACGGGATGACATGATAAATAATTCACCAGAA 300
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGln--- 115
Db 301 GATCTGAAGCAACAGTACTGACCAAGGACCAAGCAAGGAAAAAAGTGCAGAAATAAA 360
```

QY 115 ----- 115
Db 361 TATGGCATGACTAAAGCTTATCTTGGGGGTGCTGACATCTCTGACTCGAATAATAAACAAC 420
QY 116 ----- -GlutIleSerGlnAlaMetGluGlnGluGly 125
Db 421 AAGTATGTTGGAATTCATCTCTTTCAGAAATTTTCAGAGCTATGGAACAAGAAGGT 480
QY 126 AlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer 145
Db 481 GCCACAGCAGCAGACAGAGAAGAAATGAGGATCATGAGGTGACATGGGACTACAGAGT 540
QY 146 HisValMetThrLysPheAlaGluGluGluAspValArgArgSerPheGluAsnThrAla 165
Db 541 CACGTGATGACCAAAATTCGCTGAGAGGAGAGATGTACGTCTGTAGTTTGAAGAACACTGCT 600
QY 166 AlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPhe 185
Db 601 GCTGACTGGCCGGAAATGCAAAAGTTGGCTGGTGGCTTTTGATTCAGACCGGTGGGGCTTC 660
QY 186 ArgProArgThrValValLeuHisGlyLysSerGlyLeGlyLysSerAlaLeuAlaArg 205
Db 661 CGGCTTCGACCGTGGTCTCGACGGAAGTCAGGAATGGGAATCGGCTCTAGCCAGA 720
QY 206 ArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhe 225
Db 721 AGGATCGTGTGCTGGCGCAAGGTGGACTCTACCAAGGGAATGTCTCTCTACGCTTTC 780
QY 226 PheLeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSer 245
Db 781 TTCCTCCCGTTAGAGATGCAGCGGAGAGAGAGAGCATGTCTACAGAGTTTCACTTCC 840
QY 246 ArgGluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeu 265
Db 841 AGGAGTGGCCAGACTCCAGGCTCCGGTGCAGGAGATCATGTCCCGACAGAAAGGCTG 900
QY 266 LeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeu 285
Db 901 TTGTTCATCATGTACCGGTTTCGATGACCTGGGCTCTGTCTCCCAACAATGACACAAAGCTC 960
QY 286 CysLysAspTrpAlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLys 305
Db 961 TGCAAAAGACTGGGCTGAGAACAGCTCCGTTTCACTTACCTCATACGAGTCTGCTGAGGAAG 1020
QY 306 ValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeu 325
Db 1021 GTCTCTGCTCCCTGAGTCTCTCTCATCGTCAACCGTCAGAGACGTGGGCGACAGAGAAGCTC 1080
QY 326 LysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArg 345
Db 1081 AAGTCAGAGGTGGTGTCTCCCGGTACTCTGTAGTGTAGAGAAATCTCCGGGGAACAAAGA 1140
QY 346 IleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAla 365
Db 1141 ATCCACTTGCTCTCTGAGCGCGGATTTGGTGAGCATCAGAAGACACAAAGGGTTCGCTGCG 1200
QY 366 IleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIle 385
Db 1201 ATCATGAACAACCGTGAGCTCTCGACCGAGTGCAGGTCGCCAGGTGCCCGCGTCTCTCATC 1260
QY 386 CysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThr 405
Db 1261 TGGCTGGGCCCCGTGACGTGACGAGCTGTGGGGGAGAGCGTCGCCCTTCAACCAAAACG 1320
QY 406 LeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArg 425
Db 1321 CTCACAGGCTTGACCGCGCTTTTGTGTTCATCAGCTCACCCCTCGAGGGGTGGTCCGG 1380
QY 426 ArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGlu 445
Db 1381 CGCTGTCTCAATCTGGAGAAAGAGTGTCTCTGAAGCGCTTCTGCCGTATGGCTGTGGAG 1440

QY 446 GlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGly 465
Db 1441 GGAAGTGTGGAATAAGGAAGTCAGTGTGTGACGGGTGAGCAGACCTCATGGTTCAAGGACTCGGG 1500
QY 466 GluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGlu 485
Db 1501 GAGTCTGAGCTCGTCTGTGTTTCATAGAAACATCTTCTCCACAGACGCACTGTGAG 1560
QY 486 GluTyrTrpThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpVal 505
Db 1561 GAGTACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGTCGCGCTTGTACTACGTG 1620
QY 506 LeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArg 525
Db 1621 TTAGAGGGCCCTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACCTTGAGAAAGACAAAGAGG 1680
QY 526 SerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeu 545
Db 1681 TCCATGGAGCTTAACACAGCAGGCTTCCATATCCACTCGCTTGGATGAAGCGTTCTTG 1740
QY 546 PheGlyLeuValSerGluAspValArgProLeuGluValLeuLeuGlyCysProVal 565
Db 1741 TTTGGGCTCTGTGAGCGAAGACGTAAAGAGGCCACTGGAGGTCCTGTGGGCTGTCCGCTT 1800
QY 566 ProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsn 585
Db 1801 CCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGGTCTCTGTGGGTGACGAGCCTAAT 1860
QY 586 AlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLys 605
Db 1861 GCCACCAACCCAGGACACACCTTGGAGCGCTTCCACTGTCTTTCGAGACTCAAGACAAA 1920
QY 606 GluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsn 625
Db 1921 GAGTTTGTTCGCTTGGCATTAACAGCTTCAAGAGTGTGGCTTCCGATTAACACAGAAC 1980
QY 626 LeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArg 645
Db 1981 CTGGACTGTGATGACATCTTCTTCTGCTCCACACTGCCGTATTTGGCGAAATTCGG 2040
QY 646 ValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValPro 665
Db 2041 GTGAGTGTCAAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGCGATGTCTCTGTGGTCCCT 2100
QY 666 LeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeu 685
Db 2101 CTATGGATGGGATTAAGACCTCTCATTTGAGGAGCAGTGGGAGATTTCTGTCCATGCTT 2160
QY 686 GlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAla 705
Db 2161 GGCAACCCACCACTGCGGCGAGCTGGACCTGGGCGAGCAGCATCTTGACAGAGCGGCGCC 2220
QY 706 MetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPhe 725
Db 2221 ATGAAGACCTGTGTGCCAAGCTGAGGCATCCCACTGCCAAGATACAGACCTTGATGTTT 2280
QY 726 ArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
Db 2281 AGAAATGCACAGATTACCCCTGTGTGTGCAGCACTCTGGAGAAATCGTCAATGGCAACCGT 2340
QY 746 AsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGluAspValArgMetAla 765
Db 2341 AACCTAAGATCCCTCAACTTGGGAGGACCCCACTGAAGAGAGGATGTAAAGATGGCG 2400
QY 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGly 785
Db 2401 TGTGAAGCCTTAAACACCAACCAATGTTTGTGTGAGTCTTTGAGGCTGGATTCGTGGA 2460
QY 786 LeuThrHisAlaCysTyrTrpLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
Db 2461 TTGACCATGCTGTGTACCTGAAGATCTCCCAATCTTACGACCTCCCCCAGCCTGAAA 2520
QY 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825

```
Db      2521 TCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCAGGAGTAATGCCTCTCAGTGATGCC 2580
Qy      826 LeuArgValSerGlnCysAlaLeuGlnLysLeuLeuLeuGluAspCysGlyIleThrAla 845
Db      2581 TTGAGAGTCTCCAGTGGCCCTGCAAGAGCTGATCTGAGGAGCTGTGGCATCACGCC 2640
Qy      846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
Db      2641 ACGGGTTGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTGC 2700
Qy      866 LeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuLeuCysArgSerMetArgLeu 885
Db      2701 CTATCCAAACACAGCCTGGGAAACAAGGTGTAATCTACTGTGTGATCCATGAGGCTT 2760
Qy      886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
Db      2761 CCCCACTAGTCTGCAGAGGCTGATGCTGATCATGTCACCTGGACACCGCTGGCTGT 2820
Qy      906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMet 925
Db      2821 GGTTCCTTTCACCTTGGCTTATGGGTAACTCATGGCTGACGCACCTGAGCCTTAGCATG 2880
Qy      926 AsnProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCys 945
Db      2881 AACCCCTGTGGAAGACAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACCATCTTGT 2940
Qy      946 HisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeu 965
Db      2941 CATCTCCAGGACCTGGAGTGTGTAAGTGTCTATCTACCGCGCGCTGTGTGAGAGTCTG 3000
Qy      966 SerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeu 985
Db      3001 TCCTGTGTGATCTCGAGGAGCAGACACCTGAAGACCTGGATCTCAGCGACAATGCCCTG 3060
Qy      986 GlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThr 1005
Db      3061 GGTGACGGTGGGGTGTGTCACATGTGCGAGGAGCTGAAGCAAAAAGAACAGTGTTCGAGC 3120
Qy      1006 ArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAla 1025
Db      3121 AGACTCGGGTGTGAAGCATGTGACATCTGATTCGTGTGAGGCACTCTCTCTGGCC 3180
Qy      1026 LeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAspPheSerProLys 1045
Db      3181 CTTTCCTGCAACCGGCATCTGACCACTTAAACCTGGTGCAGATAACTTCAGTCCCAAA 3240
Qy      1046 GlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeGly 1065
Db      3241 GGAATGATGAAGCTGTGTTCGGCCCTTGCCTGCCACGCTTAACCTTACAGATAATTGGG 3300
Qy      1066 LeuTrpLysTrpGlnTyProValGlnIleArgLysLeuLeuGluValGlnLeuLeu 1085
Db      3301 CTGTGGAATGGCACTACCTGTGCAATAAGAAAGCTGCTGGAGAAAGTGCAGCTACTC 3360
Qy      1086 LysProArgValIleAspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db      3361 AAGCCCGGAGTCTGAATGACGGTAGTGGCATCTTTTGTATGAAGATGACCGGTAC 3417

RESULT 10
AC011470/c
LOCUS      AC011470 157141 bp DNA linear PRI 15-JUL-2000
DEFINITION Homo sapiens chromosome 19 clone CTC-490M10, complete sequence.
ACCESSION AC011470
VERSION    AC011470.5 GI:9211204
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 157141)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
```

```
JOURNAL REFERENCE
AUTHORS 2 (bases 1 to 157141)
TITLE   Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 157141)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE   Direct Submission
JOURNAL Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT  On Jul 15, 2000 this sequence version replaced gi:7690109.
Draft sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number Of Errors is 0.1.
```

```
FEATURES
source 1. 157141
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-490M10"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 3,87e-199 Length: 157141
Score: 2814.00 Matches: 658
Percent Similarity: 61.25% Conservative: 69
Best Local Similarity: 55.43% Mismatches: 175
Query Match: 39.78% Indels: 290
DB: 9 Gaps: 34
```

US-10-066-521-6 (1-1344) x AC011470 (1-157141)

```
Qy      135 GlyHisGlyGlyAspThrTrpAspTyLysSerHisValMetThrLysPheAlaGluGlu 154
Db      42000 GGACATGGAGGTGACACATGGGACTACAGAGTCACGTGATGACCAATTCGCTGAGGAG 41941
Qy      155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db      41940 GAGGATGTACGTCGTAGTCTTTTGAACACACATGCTGCTGCTGCGCGAAATGCAACACGTTG 41881
Qy      175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
Db      41880 GCTGTGCTTTTGATTGACACCGGTGGGGCTTCGCGCTCGACGGTGGTTCTGCACGGA 41821
Qy      195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 214
Db      41820 AAGTCAGGAATGGGAATCGGCTCTAGCCAGAGGATCGTGTGCTGGCGCGCAAGGT 41761
Qy      215 GlyLeuTyrglnGlyMetPheSerTyValPhePheLeuProValArgGluMetGlnArg 234
Db      41760 GGACTCTACAGGGAATGTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATGACAGCGG 41701
Qy      235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
Db      41700 AAGAAGGAGAGCAGTGTGCAGAGATTCTCTCCAGGAGGTGGCCAGACTCCCGAGGCTCGG 41641
Qy      255 ValThrGluIleMetSerArgProGluArgLeuPheIleLeuAspGlyPheAspAsp 274
Db      41640 GTGACGAGATCATGTCCGACCAAGAGGCTGTGTTCATCATGCGGTTTCGATGAC 41581
Qy      275 LeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db      41580 CTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTTGCAAGAGCTGGGCTGAGAACGAGCCT 41521
Qy      295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db      41520 CCGTTCCCTCATACGACGCTCTGAGGAAGGTCCTGCTCCCTGAGTCTCTCTCTCTCTCTCT 41461
```


QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
| | | | |
Db 41460 GTCACCGTCAGAGAGCTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTCTCCCGGTAC 41401
| | | | |
QY 335 LeuLeuValArgGlyLysSerGlyGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
| | | | |
Db 41400 CTGTTAGTTAGAGGAATCTCGGGGAAACAAGAATCCACTTGCTCTTGAGCGGGATT 41341
| | | | |
QY 355 GlyGluHisGlnValThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
| | | | |
Db 41340 GGTGAGCATCAGAGACACAAGGGTTGGTGGCATCATGAACAACCGTGAGCTCTCGAC 41281
| | | | |
QY 375 GlnCysGlnValProAlaValAlcYSerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394
| | | | |
Db 41280 CAGTGCAGGTGCCCGCTGGGCTCTCTCATCTGGGTGGCCCTGCAGCTGCAGAGCGTG 41221
| | | | |
QY 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
| | | | |
Db 41220 GTGGGGAGAGCGTGGCCCTTCAACCAACCGCTCAGAGCCCTGCACGGCGCTTTGTG 41161
| | | | |
QY 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgVal 434
| | | | |
Db 41160 TTTTATCAGCTCACCCCTCGAGGGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAGATT 41101
| | | | |
QY 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
| | | | |
Db 41100 GTCCTGAAGCGCTTCGCGCTATGGCTGTGGAGGGAGTGTGGATAGGAAGTCAGTGT 41041
| | | | |
QY 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
| | | | |
Db 41040 GACGCTGACGACCTCATGGTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTGCTTTTAC 40981
| | | | |
QY 475 MetAsnIleLeuLeuProAspSerHisCysGluGlnValThrPhePheHisLeuSer 494
| | | | |
Db 40980 ATGAACATCTCTTCCACAGACAGCCACTGTGAGGAGTACTACACCTCTTCCACCTCAGT 40921
| | | | |
QY 495 LeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAla 514
| | | | |
Db 40920 CTCAGGACTTCTGTGGCGCTTGTACTAGCTGTAGGGGCTTGGAAATCGAGCCAGCT 40861
| | | | |
QY 515 LeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe 534
| | | | |
Db 40860 CTCTGCCCTCTGTACGTTGAGACACAAAGAGTCCATGGAGCTTAAACAGGCAGGCTTC 40801
| | | | |
QY 535 HistLeHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
| | | | |
Db 40800 CATATCCACTCGCTTGGATGAAGCGTTCTTGTGTGGCCTCGTGAGCGAAGACGTAAGG 40741
| | | | |
QY 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
| | | | |
Db 40740 AGGCCACTGGAGGTCCTGTGGGTGTCCCGTTCCTTGGGGTGAAGCAGAGCTTCTG 40681
| | | | |
QY 575 HistTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAsp 594
| | | | |
Db 40680 CACTGGGTCTCTCTGTTGGGTGACGAGCTAATGCCACCCAGGAGACACCCCTGAC 40621
| | | | |
QY 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
| | | | |
Db 40620 GCCTTCCACTGCTCTTTTCGAGACTCAAGACAAAGAGTTGTTCGCTTGGCATTAACAGC 40561
| | | | |
QY 615 PheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCys 634
| | | | |
Db 40560 TTCCAGAGAGTGGCTTCCGATTAAACAGNACCTGGACTGTATAGCATTTCTCTTCG 40501
| | | | |
QY 635 LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg 654
| | | | |
Db 40500 CTCAGCACGTCTCGGTATTTGCGGAAATTCGGGTGGATGTCAAGAGGATCTTCCCAAGA 40441
| | | | |
QY 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667
| | | | |
Db 40440 GATGAGTCCGCTGAGGCATGCTCTGTGGTCCCTCTATGGTGAGTACCCAGGCGAGTTT 40381
| | | | |
QY 667 ----- 667

Db 40380 TCCTATGCCCGTGTCTGCTGAGCTCTGTCTCTACTCTGCTGGGACTTGACATGACTTCCAGGA 40321
| | | | |
QY 668 ---MetArgAspLysThrLeuIleGlu-----GluGlnTrp--- 678
| | | | |
Db 40320 ACTTCAAGGTCCAGAGAAATCTTTCCAGGATGAATGGCCAGATGAGCTCCAGCTGGCTA 40261
| | | | |
QY 679 -----GluAsp-----PheCysSerMetLeuGly-----Thr 687
| | | | |
Db 40260 AAATGCCAGGACCGGAATCTGGCTCATTTGTCACACTGGGAAAGGTCCTTTCAATACA 40201
| | | | |
QY 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArg----- 704
| | | | |
Db 40200 GGGCCCGCCCTAGA-----CTAAAGGAAGGCTCTGGGT 40165
| | | | |
QY 705 -----AlaMet 706
| | | | |
Db 40164 TGAACCCAGGAATATGTGGTATTTATTTTAAAGTCAGAGTCTCACTCTGTGGC 40105
| | | | |
QY 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
| | | | |
Db 40104 AGCTGGAGGTAGTGGCAGCATCTCGGTCACTGCCAACTCCACCTCCAGGCTCAAGA 40045
| | | | |
QY 727 AsnAlaGlnIle-----ThrProGlyValGlnHis----- 736
| | | | |
Db 40044 GATCTCTCCCTCAGCTGCCGAGGAGCTGAGGACTACAGGCAATGCACCAATGCCAG 39985
| | | | |
QY 737 LeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHis 756
| | | | |
Db 39984 CTAGTTTTGTAAATTTTAGTAGAGATGGGATTTTCG-----CACGTTGGCCAGGCTGGT 39931
| | | | |
QY 757 LeuLysGluLeuAspValArg----- 763
| | | | |
Db 39930 CTCGAAT -CTGACCTCAGGTGATCCACCGCTCAGCCTCCCAAAGTCTGGGATTAC 39872
| | | | |
QY 764 -----MetAlaCysGluAlaLeu----- 769
| | | | |
Db 39871 AGGCATCAGCAGCCCTCCCATGATGTGACCGCTCAGCGTTGTGGACTGTATTATGG 39812
| | | | |
QY 770 -----LysHisProLysCys 774
| | | | |
Db 39811 TATGTTGTCAGGGCGGTGTCTGTCCAGTAGAGGGTTCTGTGATGGAATATTCATGTGC 39752
| | | | |
QY 775 Leu-----LeuGluSerLeuArgLeu----- 781
| | | | |
Db 39751 CTGTCCAGTAGACAGCCACCCAGCCCTCATAGCCCTGTTGAGCATCTGAACATGTGTAG 39692
| | | | |
QY 782 ---AspCysCysGlyLeuThrHis-----AlaCysTyrLeu 792
| | | | |
Db 39691 TGAGACTGTGGAGTGAAACTCATCTCACTCCATTTCAACTCATTTGTATTTT-TTTCYT 39633
| | | | |
QY 793 LysIleSerGlnIleLeuThrThrSerProSerLeuLys----- 805
| | | | |
Db 39632 TTTTTCAGACAGAGTCTCACTGTGTCCACCAAGCTGGAGTACAGCAGCATGATCTTTGCT 39573
| | | | |
QY 806 -----SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSer 823
| | | | |
Db 39572 CACTGCAATCTCTGCCCTCCCGGGTTCA-----AGCGATTCTCTCGA-CTCAGCCTCTCA 39520
| | | | |
QY 824 AspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle-----LeuGluAsp 840
| | | | |
Db 39519 AGTAGCTGGGATATAGGCAGTGGCCACCATCTGCTGCTAAATTTTGTATTTTAGTAGAG 39460
| | | | |
QY 841 Cys-GlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg-- 859
| | | | |
Db 39459 ACAGGGTTTCAAC-----ATGTTGGCCAGGCTGATCTCGAATCTCCTG 39418
| | | | |
QY 860 ----SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLe 878
| | | | |
Db 39417 ACCTCAGGTGAGCCACTGCC-TTGGCCTCCCAAGTGTGGGATTTATAGGCATAGCTG 39359
| | | | |
QY 878 uLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArg----- 892
| | | | |

Dbb 39358 CTGTGCCCGG-----CCTCATTTGTAATTAATAGTCACACATAGCTGATGCT 39311
Qy 893 -----LeuMetLeuAenGlnCysHis----- 899
Db 39310 TGCTGAATAGTCAGGCTAGAAAACACACTTGTCTGA-CTGTCTAGTGTCTATCTCC 39252
Qy 900 -----LeuAepThrAla-----GlyCysGlyPheLeuAl 909
Db 39251 TATGACCCAGAACACAGCAGTCAGCAATGTGCAGAGCTGGTGGACGAGGATTACTTGA 39192
Qy 909 aleuAlaLeuMetGlyAenSerTrpLeuThr-HisLeuSerLeuSerMet----- 925
Db 39191 GGCAGGTTTTTGTAGACAGCAGCTGGGCAACAAAGTAAGACCTGTCTGTGTGACAGGCGG 39132
Qy 926 -----AenPro-----ValGluA 930
Db 39131 GTGCAGTGCATCATCTGTGTAATCCAGCAGCTTTGGGAGGCTGAGTGGTGCATCAGA 39072
Qy 930 spAenGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAepL 950
Db 39071 GGTGAGGATTCGA-----GACCAGCTGGGCCAACATAGTGAAACCC 39030
Qy 950 euGluLeuValLys-----CysHisLeuThrAlaAlaC 961
Db 39029 CATCTCTATAAAATACAAAATTAGCCGATGTGTGGCGCACACCTGTAATCCGAGC 38970
Qy 961 ysCysGluSerLeu-----SerCysValLies 970
Db 38969 TACTCAGACGCTTGAGGCGAGGAGAAATTCCTGAACCCAGGAGGAGGTTGCAGTGAGC 38910
Qy 970 erArgSerArgHisLeuLysSerLeuAepLeuThrAspAsnAlaLeuGlyAepGlyGlyV 990
Db 38909 CGAGATCAGCCGCTGCTCCAGCTGCGGCAACAGAG-----T 38871
Qy 990 alAlaAlaLeuCysGluGlyLeuLysGlnLysAenSerValLeuThrArgLeuGlyLeuL 1010
Db 38870 GAGACTCCATCAAAAAATAAAAAATAAAAAAGCCCTCAAAATTATCACATGGGGCTGT- 38812
Qy 1010 ysAlaCysGlyLeuThrSerAepCysCysGluAlaLeuSerLeuAla----- 1025
Db 38811 -----GGCGCGTGTCTATGCTGTATCTTAGCAGCTTTGGGAGGCTGAGCGAGGT 38760
Qy 1026 --LeuSerCysAsnArgHisLeu-----ThrSerLeuAenLeuValGln-AsnAsnPheSer 1043
Db 38759 GGATCACACTAGTTCAGGAGTTCGAGACCGCTGGCCCAACATAGTGAACCCCATCTCT 38700
Qy 1044 ProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIle 1063
Db 38699 GCTAAA-----AATACAAAAT 38682
Qy 1064 IleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeuGluValGln 1083
Db 38681 AGTGGGGGT 38622
Qy 1084 LeuLeuLysProArg 1088
Db 38621 TTGCTGGAACCCCGC 38607

RESULT 11
AC024580
LOCUS AC024580 193609 bp DNA linear PRI 21-DEC-2001
DEFINITION Homo sapiens chromosome 19 clone CTD-2621117, complete sequence.
ACCESSION AC024580
VERSION AC024580.6 GI:17975240
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 193609)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 193609)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (13-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 193609)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Dec 21, 2001 this sequence version replaced gi:16905144.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.9.

FEATURES
source
1. 193609
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2621117"

ORIGIN

Alignment Scores:
Pred. No.: 5,25e-199 Length: 193609
Score: 2814.00 Matches: 533
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.78% Indels: 0
DB: Gaps: 0

US-10-066-521-6 (1-1344) x AC024580 (1-193609)

Qy 135 GlyHisGlyLeuGlyAepThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 154
Db 22258 GGACATGGAGTGCACATCGGACTACAGAGTCACGATGATGACCAATTCGCTGAGGAG 22317
Qy 155 GluAspValArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 22318 GAGGATGATCGTGTAGTTTGAACACACTGCTGCTGACGCGGAAATGCAAAACGTTG 22377
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
Db 22378 GCTGCTGCTTTGATTCAGACCGGTGGGGCTTCGGGCTTCGACGGTGTTCGACGGA 22437
Qy 195 LysSerGlyLeuGlyLysSerAlaLeuAlaArgArgLysValLeuCysTrpAlaGlnGly 214
Db 22438 AAGTCAGGAATTCGGAATTCGGCTCTAGCCAGAGGATCGTGTGCTGCTGGGCGCAAGGT 22497
Qy 215 GlyLeuTrpGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
Db 22498 GGACTCTACAGGGAATGTCT 22557
Qy 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
Db 22558 AAGNAGAGAGCAGTGTTCACAGATTTCATCTCCAGGGAGTGGCCAGACTCCAGGCTCCG 22617
Qy 255 ValThrGluLeuMetSerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspAsp 274
Db 22618 GTGACGAGATCATGTCCGACACAGAAAGCGTGTGTGTTCATTCATTCATTCATTCATTCAT 22677
Qy 275 LeuGlySerValLeuAenAenAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 22678 CTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGCAAAGACTCTGGGCTGAGAACGCTCT 22737

QY 295 ProPheThrLeuLeuAArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuLeu 314
DB 22738 CCGTTACCCCTCATACGCGAGTCTGCTGAGGAGGTCTCTGCTCCCTGAGTCTCTCTGATC 22797
QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
DB 22798 GTCACCGTCAGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTCGTGTCTCCCGGTAC 22857
QY 335 LeuLeuValArgGlyLysSerGlyGluClnArgLysHisLeuLeuLeuGluArgGlyLe 354
DB 22858 CTGTAGTTAGAGGAATCTCCGGGGAAACAAGATCCACTTGTCTCTGAGCGGGATT 22917
QY 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
DB 22918 GGTGAGCATCAGAGACACAGGGTGGCTCGATCATGAACAACCGTGAGCTGCTCGAC 22977
QY 375 GlnCysGlnValProAlaValGlySerLeuLeuLeuValAlaLeuGlnLeuGlnAspVal 394
DB 22978 CAGTGCCAGGTGCCCGCGTGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTG 23037
QY 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
DB 23038 GTGGGGGAGAGCGTGGCCCCCTTCAACCAACACGCTCACAGCCCTGCAGCGCGCTTTTGTG 23097
QY 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgVal 434
DB 23098 TTTTATCAGCTCACCCCTCAGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGAAAGATT 23157
QY 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
DB 23158 GTCTGAAGCGCTTCTGCGGTATGGCTGTGGAGGAGTGTGGATATGAAGTCAAGTGTGT 23217
QY 455 AspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
DB 23218 GACGGTCACGACCTCATGGTTCAGAGGACTCGGGAGTCTGAGCTCCGCTGCTGTTCAC 23277
QY 475 MetAsnLeuLeuProAspSerHisCysGluGluTyrTrpPhePheHisLeuSer 494
DB 23278 ATGAACATCTCTTCCACAGACGACCTGTGAGAGTACTACCTTCTTCCACCTCAGT 23337
QY 495 LeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluLeuProAla 514
DB 23338 CTCCAGGACTTCTGTGCGCCTGTACTACGTGTAGAGGCGCTGGAATCGAGCCAGCT 23397
QY 515 LeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe 534
DB 23398 CTCTGCCCTCTGTACGTGTGAAGACAAAGAGGTCATGAGGCTTAAACAGGCGGCTTC 23457
QY 535 HisIleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
DB 23458 CATATCCACTCGCTTGTGATGAAGCGTTTCTGTGTGGCTCTGTGAGCGAAGACGTAAG 23517
QY 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
DB 23518 AGGCCACTGGAGGTCCTGCTGGGCTGCCCTTCCCTCGGGGTGAAGAGCTTCG 23577
QY 575 HisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAsp 594
DB 23578 CACTGGGTCTCTCTGTGGGTTCAGACGCTAAATGCCACACCCCGAGAGACACCTCGAC 23637
QY 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
DB 23638 GCCTTCCACTGCTTTTCGAGACTCAAGACAAAGAGTTGTTCGCTTGGCATTAACAGC 23697
QY 615 PheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCys 634
DB 23698 TTCAAGAAAGTGGCTCCGATTAAACAGAACCTGGACTGTATAGCATCTTCTCTCTCG 23757
QY 635 LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg 654
DB 23758 CTCAGACACTGTCCGTATTTTCGGGAAATTCGGGTGGATGTCAAGAGGATCTTCCCAAGA 23817
QY 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667

DB 23818 GATGAGTCCGCTGAGGATGCTCTGTGTCCCTCTATGG 23856
RESULT 12
AY196362 3470 bp mRNA linear ROD 18-DEC-2003
LOCUS Mus musculus strain CBA/J WATER protein (Water) mRNA, complete cds.
DEFINITION AY196362
ACCESSION AY196362 GI:37781353
VERSION AY196362.1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3470)
AUTHORS Cheng,H., Gao,J., Zhang,Y., Huang,Z., Liu,H., Teuscher,C. and Ma,R.Z.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-2002) Human and Animal Genetics Division,
Institute of Genetics & Developmental Biology, Chinese Academy of
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China
FEATURES
Location/Qualifiers
1..3470
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CBA/J"
/db_xref="taxon:10090"
/chromosome="7"
1..3470
/gene="Water"
9..3344
/gene="Water"
/codon_start=1
/product="WATER protein"
/protein_id="AA052698.1"
/db_xref="GI:37781354"
/translation="MGPPKSKAILKARGLEEEKSEKMTSPENDSKSIQKDGPE
QQTSESTGPPPEKSKAILKARGLEEEKSESTSPSENKAILKSGSREVEQAS
ERKMTSPENDSKSIQKDGPEQSTSETLQSEBEDEVTEADKNGDGLQDYKAHVIAK
RFTSDVLDHYDSPEMKLLSDAFPYQKTFQPHIILHGFVGKGSALARSIVLGAQK
LQKQSFVIFSVRIKWTKEKSLAQIAKECPDSDPVTKIMSPERILLFVLDGDD
MDFQHDQMTLSRDWKDQPIYILMYSLRKLALPOSFLITTRTGLKLSMVVS
PLYLVGSLASRRSOLVLENTSNESDRIOVFHSLIENHOLFQCOAPSVCSVCEAL
QOKLKGKCTLPCTLTGLYATLVFHMILTKRPSQALSQEQITLVGLCMAAEV
WTMSVYFDDDLKNSYKSEILALFHMILQVHNSQCYFVSHLSQDFFAALYY
VLEGEEWQHFCFTENQSRIMEVKTDTLLGMRKFLGLMKNKDLKTLVLEFPY
VTPTEQLQHWVSLTAQQVNGTSPMDTLDAFYCLFESQDEBFVGALXRFVWLLI
NOKMDLVSSYCLKHCONLKAIRVDIRLLSDVNTLELCPVTVQTCCKPILMBWVG
NFCSVLSGLNLEKLDIGSLISQAMKILCLELRNQCRIQKLTFKSAEVVSGLKH
WKLFSNQMLKYNLGNTPMKDDMKLACEALKHPKCSVETLRLOSCELTIIGYEMIS
TLLISTTRUKLSLAKNRVGVKSMISLGNALSSMLQKILNDNGCLTPASCHLVS
ALFNSQNLTHLCLSNNGTEGVQQLCFLRNPECALQRLILHNKILVDDAYGFLAMR
LANNTKLHLSLTWNPVGDMKLCEALKEPTCYLOEVLVDCOLTQNCEDLACMI
TTTKOLKSLDIGNALGDGVTLCCEGLKQNNSSLRRLGLGACKLTNSCCALSLAIS
CNPLNSLNLVQNDFTSGMLKLSAFQCPVSNLGTIGLWKQEVYARVRQLEVEFV
KPHVIDGWYASDEDDRNWKN"

ORIGIN

Alignment Scores:
Pred. No.: 1,36e-192 Length: 3470
Score: 2695.00 Matches: 555
Percent Similarity: 66.49% Conservative: 185
Best Local Similarity: 49.87% Mismatches: 293
Query Match: 38.10% Indels: 81
DB: 10 Gaps: 11
US-10-066-521-6 (1-1344) x AY196362 (1-3470)
QY 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPheLysGluLeuLys 34
DB 163 CAATGGGTCTCCCA-GAAAAAGACAGTAAAGCAATCTTCAAAGCAGCTGGATTGGAAGAG 221
QY 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54

Db 222 GACAGAAAGTCAGAAAGCACA----- 242
Qy 55 ValGluCysLeuAlaLeuLeuHisGluTyrTyrglyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
Qy 75 SerIleSerIlePheGluAenMetAsnLeuArgThrLeuSerGluLysAlaArgAsp 94
Db 243 ---ATGTCTCTTCAGAAAATGTCTAGTAAAGCCATCTCTGAAAGACAGTGGATCAGAGAA 299
Qy 95 MetLysLysIleSerGln----- 100
Db 300 GTGGAAACAGGCATCAGAAAGAAAATGACTTCTCCAGAAAACGACAGTAATCAATCCAG 359
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGln 115
Db 360 AAAGACCAAGCAGCAGCAGGAGCAGACATCAGAAACCTTACATCTAAGGAAGAGAT 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135
Db 420 GAAGTCAGAGAGCA-----GATAAAGAT 443
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155
Db 444 AATGAGGTGACTTACAGAGCTACAAAGGCCATGTGATTGTTAAAGTTCGACACAAGTGTG 503
Qy 156 AspValArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCAGAGATGAATATTGTCT 542
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 543 GATGCTTTTAAACCATACAGAAAACCTTCCAGCCTCACACCATTTATCTACATGGAAGA 602
Qy 196 SerGlyIleGlySerAlaLeuAlaArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 603 CCAGGAGTTGGAGTACGCTTGGCCAGAGATATTGTTTGGCTGGGCACAGGGTAAA 662
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
Db 663 CTCCTCCAAAAGTG---TCCTTGTGCATCTCTCTCTGTAGAGAAATAAAGTGACA 719
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 720 GAGAAGACAGCTTGGCACAGCTGATTCTAAGGAGTGTCCAGACTCTCTGGGATCCAGTG 779
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspLeu 275
Db 780 ACAAGATCATGTCCCAACCAAGAAAGACTCTTGTGTGTCATAGATGGCTTGGATGATG 839
Qy 276 GlySerValLeu---AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 840 GACTCTGCTCCAAACATGATGATATGACATATCCAGAGACTGGAAGATGAACAGCCC 899
Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 900 ATATACATCTGATGATCAGACCTCTCTGAGGAAGGCTCTCTTACCTCAGTCTCTTCTCATC 959
Qy 315 ValThrValArgAspValGlyThrGluLysLysSerGluValValSerProArgTyr 334
Db 960 ATTACCACGAGAAACACAGGCTTAGAAAACCTCAAGTCAATGTTGTTTCCCTCTCTAT 1019
Qy 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 1020 ATACTGTTGAGGACTGTCTGCATCAGGAGAGATCTCAGCTGGTCTCTGAGAAACATCTCC 1079
Qy 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuAsp 374
Db 1080 AATGAGTCTGATAGATAAAGTCTTCCATCTCTCTGATAGAAAATCACCAGCTGTTTGAC 1139
Qy 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394

Db 1140 CAATGCCAGGCCCTCTGTGTGTCTCCTCGTCTGTGAGGCTCTGCAGCTACAGAGAAA 1199
Qy 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
Db 1200 CTGGAAAGAGATGACACCTACCTACCTGCCAGACTCTCACCGGTTGTATGCCACGTTGGTG 1259
Qy 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAenLeuGluGluArgVal 434
Db 1260 TTTCCACAGCTCACCTTGAAGAAGCCCTTCCAGAGCGCTCTCAGTCAGGAAGAAGACAGATT 1319
Qy 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
Db 1320 ACTTAGTGGTGTGTCATGATGCGAGCTGAAGAGTGTGGACCATGAGGTCCGTTGTC 1379
Qy 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
Db 1380 TATGATGATGATCTGAAGAACTATAGCTTAAAGAGTCTGAGATCTTGGCCCTCTTTCAC 1439
Qy 475 MetAsnIleLeuLeuProAspSerHis---CysGluGluTyrTyrThrPhePheHisLeu 493
Db 1440 ATGAACATCTCTCCAGGTGGCCACAACAGTGCAGTGTATTGTTTCTCCACCTC 1499
Qy 494 SerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeu---GluLeuGlu 512
Db 1500 AGCTGCAGGATTTCTTGTGCTTATATATGTTTTAGAAGGCTGGAGGAATGGAAT 1559
Qy 513 ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla 532
Db 1560 CAGCATTTTTC-----TTCATTGAAAACCAAGGAGCATCTGGAGGTGAAGAACT 1613
Qy 533 GlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAsp 552
Db 1614 GAC---GACACTCGCCTCTTGGGATGAAGGTTTCTTATTTGGCCTCATCAACAAGAT 1670
Qy 553 ValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLys 572
Db 1671 ATCTTGAAGACTCTGGAGGTTCTGTTTGAATATCCCGTGAATCCCACTGTTGAGCAGAAG 1730
Qy 573 LeuLeuHisTrpValSerLeuLeuGlnProAsnAlaThrThrProGlyAspThr 592
Db 1731 CTCCAACACTGGGTCTCTCTGATAGCTCAGCAGGTCAATGGCACCCCAATGGACACC 1790
Qy 593 LeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeu 612
Db 1791 CTGATGCTCTTATTGCTATTGAGTCTCAGGATGAAGAGTTTGTGGGGGCTCTC 1850
Qy 613 AsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAsnSerSer 632
Db 1851 AAACGCTTCCAAAGATGTGGCTGCTGATTAAACCAAGATGGACTTGAAGGTCTCTTCC 1910
Qy 633 PheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePhe 652
Db 1911 TACTGTCTCAAGCACTGTGAGAACTTGAGGCAATCCGGTGGATATCAGAGACTCTCTC 1970
Qy 653 ProArgAspGluSerAlaGluAlaCysProValValProLeuTrp---MetArgAspLys 671
Db 1971 TCGGTAGATAATACTCTCGAGCTGTGCCCTGTGTTACTCTCCAGGAGACACAATGTAAAG 2030
Qy 672 ThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeu 691
Db 2031 CCCCTCTCATGAGTGGTGGGAAACTTCTGTCTGTGTGGCAGCCTCCGGAACCTG 2090
Qy 692 ArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAla 711
Db 2091 AAGAGCTGAGACTTGGCGGACAGCATCTGAGTCAACGGGCCATGAAGATACTGTGCTC 2150
Qy 712 LysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThr 731
Db 2151 GAGCTGGGAATCAGTCTCTCAGAAATACAGAAGCTGACGTTTAAAGAGTCAGAGTAGTG 2210
Qy 732 ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsn 751
Db 2211 TCTGGCCTGAACATCTCTGGAAGCTCCTTTTGAACAATCAAAACTTAAAGTACCTCAAT 2270

```
Qy 752 LeuGlyGlyThrHisLeuLeuGluGluaspValargMetAlaCysGluAlaLeuLeuHis 771
Db 2271 CTAGGAAACATCCATGAAAGATGATGACATGAAGTTAGCTCGGAAGCGCTGAAACAT 2330
Qy 772 ProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyr 791
Db 2331 CCAAGTGCTCGTGAGACTCTGAGTTGATTCCTGTGAGTTAACCATCATTTGGCTAT 2390
Qy 792 LeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGly 811
Db 2391 GAGATGATCTCCAGCTCTTATTTCACACACAGGCTAAAGTGTCTCAGCGTCCGCAAA 2450
Qy 812 AsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCys 831
Db 2451 AATAGAGGGAGTAAAGACATGATATCCCTGGGAATGCTTGAAGTAGCTCAATGTGT 2510
Qy 832 AlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeu 851
Db 2511 CTACTGCAAAAGTTGATCTGACCACTGTGGCTCTACACTGCCAGCTGCCACCTCTG 2570
Qy 852 AlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeu 871
Db 2571 GTCTCAGCCCTTTTCAGCAACACCAACTTGACACACCTGTGCTGTCTCAACACAGCCTG 2630
Qy 872 GlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGln 891
Db 2631 GGGACTGAAGAGTGCAACAGCTGTGTCAGTCTCTGAGGAATCAGAAATGCTCTCAG 2690
Qy 892 ArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAla 911
Db 2691 CGGCTGATCTGATCACTCAACATTTAGATGATGCTTATGCTTCTCTGGCAATCAGA 2750
Qy 912 LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsn 931
Db 2751 CTTGCAAAACACAAAGCTGACCCACCTGAGCCTGACCATGAAACCCGCTAGGGGATGGT 2810
Qy 932 GlyValLysLeuLeuGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGCTACTGTGTGAAGCTTTTAAAGAACTTACTTGTATACCTCAAGAACTGAA 2870
Qy 952 LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db 2871 CTAGTGAGCTGCCAACTCACACAGAACTGCTCGCAGGACCTGGCTGTATGATCACACA 2930
Qy 972 SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAAGCAATTTAAAGTTTGATCTTGGTAAACACGCCCTGGGTGACAAAGGAGTCATA 2990
Qy 992 AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGACTGAGCAAAATAACAGCTCCCTGAGGAGACTTGGTGTGGGGCA 3050
Qy 1012 CysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAGTTGACTTCAATGCTGTGAGGACTTGTTCATTGGCCATCTCTTGCACCCCTCAC 3110
Qy 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCys 1051
Db 3111 CTGAACAGCCTAAACCTGGTGAAGAAATGACTTTCAGTACATCGGGGATGTTGAAGCTGTC 3170
Qy 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyr 1071
Db 3171 TCTCGTTCANAGCTCTGCTCTTAACCTGGGATAATGCGCTGTGGAAGCAGGAGTAC 3230
Qy 1072 ProValGlnIleArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIle 1091
Db 3231 TATGCCGAGTGAGACAGACAGCTGAGGAAGTTGAGTTTGTCAAGCCCCAGCTGTGATT 3290
Qy 1092 AspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3291 GATGGTATTGGTATGTAGTGATGAAGATGACCCGAAAC 3329
```

```
RESULT 13
AY329487
LOCUS AY329487 3405 bp mRNA linear ROD 18-DEC-2003
DEFINITION Mus musculus strain CBA/J MATER protein isoform-E (Mater) mRNA,
complete cds; alternatively spliced.
ACCESSION AY329487
VERSION AY329487.1 GI:37595472
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
```

```
REFERENCE
1 (bases 1 to 3405)
Cheng,H., Zhang,Y., Huang,Z. and Ma,R.Z.
Identification and characterization of alternative splice variants
for murine Mater
```

```
UNPUBLISHED
2 (bases 1 to 3405)
Cheng,H., Huang,Z., Zhang,Y., Liu,H., Teuscher,C. and Ma,R.Z.
Direct Submission
Submitted (20-JUN-2003) Human and Animal Genetics Division,
Institute of Genetics and Developmental Biology, Chinese Academy of
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China
Location/Qualifiers
```

```
1. .3405
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CBA/J"
/db_xref="taxon:10090"
1. .3405
/genes="Mater"
9. .3296
/genes="Mater"
/notes="alternatively spliced"
/codon_start=1
/product="MATER protein isoform-E"
/protein_id="AAO94610.1"
/db_xref="GI:37595473"
/translation="MGPPKDKSKAILKARGLBEEQSKSRMTSPENDSKSIQKQGP
```

```
EQBSTSTWGPPEKSKAILKARGLBEEQSKSRMTSPENSKAILKDSGSEVEQAS
ERMTSPENDSKSIQKQGPQEQSDNGDLQDYKAHVIAKFDTSVDLHYDSPEMKL
LSDAFYKQKTFQPHITILHGRPGVGKSAARSIVLGWAQGLFQKSFVIFPSVREI
KTEKSLAQLIAKECPDSWDVPVKIMSOPERILLFVIDGLDDMDSVLQHDMDLSRDW
KQOPIYILMYSILRKALLPQSPLIITTTGTGLEKLSMVVSPLYILVGLSASRBSQ
LVLENSNEDRIQVPHSLIENHQLFDQCAQSVCSLVEALQQLKAKRCTLPCQT
LTYGLATLVFHLQTLKRPQSALSSQEEITLVGLCMAAEGVMTMSVFDLDDKNYS
LKSEIETALFHMNILLQHNSEQCYSFSLSDQFFAALYYVLEGLMEHQFCFIE
NORSIMEVKRTDTRLLGMKRFILGLMNDILKTLVLEFPVPTVEQKLQHWVLI
AOQVNTSPMDTLDAPYCLPESODEFVGALKRFOEVLWLLNOKMDLVKSSYCLHLI
QNLKATRVDIRLLSVNDTLELCPVTVQETOCKPLLMWGNFCSVLGSLNKLKLD
LGDLSILQAMKILCLELRNQSCRIOKLPKSAEVVSGLKLWKLFLFSQNLKYLNLG
NTPMKDDDKLKEALHPKSVETURLDSCELTIIGYEMIITLLITTLRLKCLSLAK
NRVGVKSMISLGNALSSCLLQKLDNCGLTTPASCHLLVSALFQNLTHLCLSN
SLGTEGVQQLCQFLRNPECALQRLILNHCNIVDDAYGFLAMRLANNTKLTSLTNP
VGDGAMKLEAKETPCYLOLELVDCOLTQNCEDLACMIITTKQLKSLDLGNAL
GDKGVTLCEGLKQNNSSLRGLGACKLTSNCELSLAIACNPHLSNLVNDPS
TSGMLKLCSAFQCPVSNLGIILGWKQEIYARVRQLEEVFVKPHVVIDGDWTASDED
DRNWWKN"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,23e-192 Length: 3405
Score: 2692.00 Matches: 557
Percent Similarity: 67.52% Conservative: 183
Best Local Similarity: 50.82% Mismatches: 294
Query Match: 38.05% Indels: 63
DB: 10 Gaps: 13
```

US-10-066-521-6 (1-1344) x AY329487 (1-3405)

```
Qy 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPhenylsGluLeuLys 34
Db 163 CAATGGTCTCTCCA-GAAAAAGACAGTAAGCAATCTTGAAAGCAGCTGGATTGGAAGAG 221
```

QY 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54
: : : : :
222 GAACAGAAAGTCAGAAAGCACA----- 242

QY 55 ValGluCysLeuAlaLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
: : : : :
242 ----- 242

QY 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
: : : : :
243 ---ATGTCCTCTCAGAAATGTGAGTAAAGCCATCTCGAAGACAGCTGGATCAGAAGAA 299

QY 95 MetLysLysIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlu 114
: : : : :
300 -----GTGGAACAGGCATCAGAAAGAAA---ATGACTTCTCCAGAAACACACAGT 347

QY 115 GlnGluIle-----SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGlu 132
: : : : :
348 AAATCAATCCAGAAAGACCAAGACCAGAGCAGGAG-----CAG 386

QY 133 GluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAla 152
: : : : :
387 ACATCAGATAATGAGGTGACTTACAGACTACAGGCCCATGTGATGCTTAAGTTGCAC 446

QY 153 GluGluGluAspValArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGln 172
: : : : :
447 ACAAGTGTGGATCTACAC-----TATGACAGCCCAAGAGATGAAA 485

QY 173 ThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeu 192
: : : : :
486 TTATTTGCTGATGCTTTTAAACCATACACAGAAACCTTCCAGCCTCACACCATATCTTA 545

QY 193 HisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAla 212
: : : : :
546 CATGGAAGACAGGAGTTGGGAAGTCAGCTTTGGCCAGAGATATGTTCTTGGCTGGGCA 605

QY 213 GlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMet 232
: : : : :
606 CAGGTAACCTCTCCAAAGAAATG---TCCTTTGTCTCTCTCTCTGTAGAGAAATA 662

QY 233 GlnArgLysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGln 252
: : : : :
663 AAGTGGACAGAGAGAGCAGTTTGGCACAGCTGATGCTGAGGAGTGTCCAGACTCTCTGG 722

QY 253 AlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPhe 272
: : : : :
723 GATCCAGTGACAAGATCATGTGCCAACCAAGAGACTCTTGTGTTGTCATAGATGGCTTG 782

QY 273 AspAspLeuGlySerValLeu---AsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291
: : : : :
783 GATGATATGGACTGTGCTCTCAACATGATGATGACACTATCCAGAGACTGGAAGAT 842

QY 292 LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
: : : : :
843 GAACAGCCCATATACATCTCTGATGATACAGCCTCTCTGAGGAAGGCTCTCTACCTCAGTCC 902

QY 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
: : : : :
903 TTTCTCATATTACACAGAAACACAGGCTTAGAAAACCTCAAGTCAATGGTGTGTTTCC 962

QY 332 ProArgTyrLeuLeuValArgIleIleSerGlyGluGlnArgIleHisLeuLeuLeuGlu 351
: : : : :
963 CCCCTCTATATCTGTTGAAGGACTGTCTGCATCAAGGAGATCTCAGCTGGTCTCTCAG 1022

QY 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
: : : : :
1023 AACATCTCCAATGAGTCTGTAGATAACAAGTCTTCCATTCTCTGTAGAAAATCACCAG 1082

QY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
: : : : :
1083 CTGTTTGACCAATGCCAGGCCCTCTGTGTGTGCTCTCTGTGAGGCTCTGCAGCTA 1142

QY 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
: : : : :
1143 CAGAAGAACTGGGAAGAGATGCACCTACCCTGCCAGACTCTCACCGGTTGTATGCC 1202

QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431
: : : : :
1203 ACGTTGGTGTTCACCAAGCTCACCTTGAAAGAGCGCTTCCAGAGCGCTCTCAGTCAGAA 1262

QY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
: : : : :
1263 GAACAGATTACTCTAGTGGGTTGTGATGATGGCAGCTGAGGAGTGTGGACCATGAGG 1322

QY 452 SerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
: : : : :
1323 TCGGTGTTCTATGATGATGATCTGAAGAACTATATGCTTAAAGGAGTCTGAGATCTTGCC 1382

QY 472 LeuPheHisMetAsnIleLeuLeuLeuProAspSerHis---CysGluGluTyrTyrThrPhe 490
: : : : :
1383 CTCTTTTCCATGATGAACATCTCTCCAGGTTGGCCACACAGTGAAGCAGTGTATGTTTTC 1442

QY 491 PheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeu--- 509
: : : : :
1443 TCCACCTCAGCTGCGAGGATTTCTTGTGCTCTATATTTATGTTTGAAGGGCTGAG 1502

QY 510 GluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeu 529
: : : : :
1503 GAATGGAATCAGCATTTTTCG-----TTCATTGAAAACCAAGGAGCATCATGGAGGTG 1556

QY 530 LysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuVal 549
: : : : :
1557 AAGAGAACTGAC---GACACTCGCCTCTTGGGATGAAGCGTTTCTTATTGTCCTCATG 1613

QY 550 SerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVal 569
: : : : :
1614 AACAGATATCTTGAAGACTCTGGAGTCTCTGTTGAATATCCCGTGATCCCACTGTT 1673

QY 570 LysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrPro 589
: : : : :
1674 GAGCAGAAGCTCCAAACACTGGGTCTCTCTGATAGTCAGAGGTCAATGGCACCAGCCCA 1733

QY 590 GlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArg 609
: : : : :
1734 ATGACACACCTGGATGCTTCTATTGTCTATTGTAGTCTCAGGATGAAGAGTTGTTGGC 1793

QY 610 LeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleLeuAsnLeuLeuAspLeuIle 629
: : : : :
1794 GGGCTCTCAACCGTCTCCAAGAGTGTGCTGCTGATTAAACCAAGAGATGGACTTGAG 1853

QY 630 AlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLys 649
: : : : :
1854 GTCTCTTCTACTGTCTCAAGCACTGTCAAGAACTTGAAGGCAATCCGGGTGGATATCAGA 1913

QY 650 GlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrp---Met 668
: : : : :
1914 GACCTCTCTCGGTAGATAATACTCTCGAGCTGTGCCCTCTGTTACTGTCTCAGGAGACA 1973

QY 669 ArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHis 688
: : : : :
1974 CAATGAAGCCCTCTCTCATGTGGAGTGTGGGAAACTTCTGCTGTGCTTGGCAGCTC 2033

QY 689 ProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThr 708
: : : : :
2034 CGGAACCTTGAGAGAGCTGGACTTGGGCGACAGCATCTGAGTCAACGGGCGCATGAAGATA 2093

QY 709 LeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAla 728
: : : : :
2094 CTGTGCTCTCGAGCTGCGGAATCAGCTCTGCAGAAATACAGAAAGCTGAGCTTTAAGAGTCA 2153

QY 729 GlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaLeuAsnArgAsnLeuArg 748
: : : : :
2154 GAGGTAGTGTGTGGCTGGAACATCTCTGGAAGCTCTCTTTTAGCAATCAAAACTTAAAG 2213

QY 749 SerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAlaCysGluAla 768
: : : : :
749 SerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAlaCysGluAla 768


```
Db 2214 TACCTCAATCTAGGGAACACCTCCCATGAAGATGATGATGAAGTTAGCTCGGAGCG 2273
Qy LeuLysHisProLysCysLeuLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHis 788
Db 2274 CTGAACATCCAAAGTGTCTGGAGACTGTGAGTTGGATCTCTGTGAGTTAACCATC 2333
Qy AlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSer 808
Db 2334 ATTGGCTATGAGATCATCTCCAGCTTCTTATTTCAACCACGAGCTTAAAGTGTCTCAGC 2393
Qy LeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgVal 828
Db 2394 CTGGCAAAATAGAGTGGGAGTAAAGCATGATATCTCTTGGGAATGCCTTGAGTAGC 2453
Qy SerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCys 848
Db 2454 TCAATGTGTCTACTGCAAAAGTTGATCTGGCAACACTGTGGCCTCACACCTGCCAGCTGC 2513
Qy GlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuLysLeuSerAsn 868
Db 2514 CACCTTCTGGTCTAGCCCTTTTCCAGCAACGAGACTTGACACACCTGTGCTGTCAAC 2573
Qy AsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCys 888
Db 2574 AACAGCCTGGGACTGAAGGAGTGCACAGCTGTGTGAGTCTCTGGAGATCCAGATGT 2633
Qy SerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeu 908
Db 2634 GCTCTCCAGCGGTGATCACTGAATCACTGCAACATTTGTAGATGATGCTTATGGCTTCTG 2693
Qy AlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProVal 928
Db 2694 GCAATGAGACTTGCACCAACACAAAGCTGACCCACTGAGCCCTGAGCAGCAACCCCGTA 2753
Qy GlnAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGln 948
Db 2754 GGGGATGTGTGCAATGAGTACTGTGTGAGCTTTTAAAGGAACCTACTTGTACTTCAA 2813
Qy AspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysVal 968
Db 2814 GAATGGAACTAGTGGAGTGGCAACTCACACAGAACTGCTGCGAGGACCTGGCCTGTATG 2873
Qy IleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGly 988
Db 2874 ATCAACAACCAAGCAATTAAGAATTGGATCTTGTATCAACAGCCCTGGGGTGACAAA 2933
Qy GlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGly 1008
Db 2934 GGAGTCATAACCTGTGTGAGGAGTGAAGCAAAATAACAGCTCCCTGAGGAGACTTGGG 2993
Qy LeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCys 1028
Db 2994 TTGGGGCATGTAAGTTGACTTCCAATTTGCTGTGAGGCAATGTCTATTTGGCCATCTTTC 3053
Qy AsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMet 1048
Db 3054 AACCTTCACTGAAACAGCCTTAAACCTGGTGAAGAAATGACTTCAGTACATCGGGATGTTG 3113
Qy LysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLys 1068
Db 3114 AAGCTGTGCTCTGGTTCATGCTCTCTAACCTGGGGATTAATTTGGCTGTGGAG 3173
Qy TrpGlnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLeuLysProArg 1088
Db 3174 CAGGAGTACTATGCCGAGTGAGAAGACAGCTGGAGGAAGTTGAGTTGTTCAGCCCCAC 3233
Qy ValValIleLeuGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3234 GTGGTATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 3281
```

```
LOCUS AX427592 3447 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from Patent WO0232955.
ACCESSION AX427592
VERSION AX427592.1 GI:21537713
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Nelson,L.M. and Tong,Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 5 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1..3447
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores: 4.54e-192 Length: 3447
Pred. No.: 2688.00 Matches: 554
Score: 66.49% Conservative: 186
Percent Similarity: 49.78% Mismatches: 293
Best Local Similarity: 38.00% Indels: 81
Query Match: Gaps: 11
DB:
US-10-066-521-6 (1-1344) x AX427592 (1-3447)
Qy 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPhelysGluLeuLeuLys 34
Db 163 CAATGGTCTCCCA-GAAAAAGACAGTAAAGCAATCTTGAAGCACCGTGGATGGAGAG 221
Qy 35 LysLysSerSerGluSerThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54
Db 222 GAACAGAGTCAGAAAGCACA----- 242
Qy 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrTyGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
Qy 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
Db 243 ---ATGTCTCTTCAAAAATGTCTAGAGCAATCTCTGAAGACAGTGGATCAGAGAA 299
Qy 95 MetLysLysIleSerGln----- 100
Db 300 GTGGAACAGCGCTCAGAAAGAAAAATGACTTCTCCAGAAAAACGACAGTAATCAATCCAG 359
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 AAAGACCAAGGACCAGCAGGAGCAGACATCAGAAACCTTACAATCTAAGGAAGAAGAT 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GAAGTCAGCAGAGCA-----GATAAAGAT 443
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 444 AATGGAGGTGACTTACAGACTCAAGGCCCATGTGATTCCTAAGTTCGACACAAGTGTG 503
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCGACAGATGAAATTTATTGTCT 542
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 543 GATGCTTTTAAACCATATCCAGAAAAACCTTCCAGCCCTCACACCATTTATCTACATGGAAGA 602
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgGlyIleValLeuCysTrpAlaGlnGlyGly 215
```

Db	603	CGAGGAGTTGGGAAAGTCAGCTTTGGCCAGAAAGTATTCTTTGGCTGGGCACAGGGTAA	662
Qy	216	LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys	235
Db	663	CTCTCCAAAAATG---TCTTTGTGCATCTTCTCTCTGTAGAGAAATAAAGTGGACA	719
Qy	236	LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal	255
Db	720	GAGAAGACGAGTTTGGCACACGCTGATTGCTAAGGAGGTGTCAGACTCCTCGGATCTAGTG	779
Qy	256	ThrGluIleMetSerArgProGluuArgLeuLeuPheIleIleAspGlyPheAspAspLeu	275
Db	780	ACAAAGATCATGTCCCAACACAGAAAGACTCTGTGTTCATAGATGGCTTGGAGATGATAT	839
Qy	276	GlySerValLeu---AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro	294
Db	840	GACTCTCTCTCCAAACATGATGATACACTATCCAGAGACTTGGAGGATGAACAGCCC	899
Qy	295	ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle	314
Db	900	ATATACATCTGTATGTACAGCTCCTCAGGAAGGCTCTCTTACCTCAGTCTCTTTCATC	959
Qy	315	ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr	334
Db	960	ATTACCAACCAAGAACACAGGCTTAGAAAAACTCAAGTCAATGGTTGTGTCCCCCTCAT	1019
Qy	335	LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle	354
Db	1020	ATACTGTTGAAGGACTGTCTGCATCAAGGAGATCTCAGCTGGTCTCTCAGAACATCTCC	1079
Qy	355	GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp	374
Db	1080	AATGAGTCTGATAGAAATACAAGTCTTCCATTCTCTAGTAAAAATCACCAGCTGTTTGAC	1139
Qy	375	GlnCysGlnValProAlaValGlySerLeuIleCysValAlaIleuGlnLeuGlnAspVal	394
Db	1140	CAATGCGAGGCCCTCTGTGTGCTCCTCGTGTGAGGCTCTTACAGCTACAGAAAGAA	1199
Qy	395	ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal	414
Db	1200	CTGGGAAGAGATGCACCTTACCCTGCCAGACTCTCACCGGTTTGTATGCCAGTGGTG	1259
Qy	415	PheHisGlnLeuThrProArgGlyValValArgCysLeuAsnLeuGluArgVal	434
Db	1260	TTTCACCAGCTCACCTTGAAAGGCCCTCCAGAGCGCTCTCAGTCAGGAAGAACAGATT	1319
Qy	435	ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe	454
Db	1320	ACTCTAGGGTTTGTGCATGATGGCAGCTGAAGGAGTGGACCATGAGCGGTGGTTC	1379
Qy	455	AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis	474
Db	1380	TATGATGATGACCTGAAGAACTATAGCTTAAAGGAGTCTGAGATCTGGGCCCTTTTCC	1439
Qy	475	MetAsnIleLeuLeuProAspSerHis---CysGluGlyTyrThrPhePheHisLeu	493
Db	1440	ATGAACATCTCTCTCCAGTTGGGCCACACAGAGTGAAGGAGTGTATGTTTCTCCACCTC	1499
Qy	494	SerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeu---GluIleGlu	512
Db	1500	AGCCTGAGGATTTCTTTGTGCTCTATATATGTTTTTGAAGGGCTTGGAGGAATGGAAT	1559
Qy	513	ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla	532
Db	1560	CAGCATTTTGC-----TTCATTGAAAAACCAAGGAGCATCATGGAGGTGAAGAGACT	1613
Qy	533	GlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAsp	552
Db	1614	GAC---GACACTCGCCTCTCGGATGAAGCGTTTCTATTGTCCTCATGAAACAAGGAT	1670
Qy	553	ValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLys	572
Db	1671	ATCTTGAAGACTCTGGAGGTTCTGTTTGAATATCCCGTGATTCACACTGTTGAGCAGAG	1730

Qy	573	LeuLeuHisTrpValSerLeuLeuGlnGlnProAlaThrThrProGlyAspThr	592
Db	1731	CTCAACACATGGGGTCTCTGTGATAGCTCAGCAGGTCAAATGGCACAGCCCAATGACACCC	1790
Qy	593	LeuAspAlaPheHisCysValLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeu	612
Db	1791	CTGGATGCCCTTCATTGTCTATTGTAGCTCTCAGATGAAGAGTTGTGTGGCGGGGCTCTC	1850
Qy	613	AsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSer	632
Db	1851	AAACGGCTTCAACGAAGTGTGGCTGCTGATTAAACACGAGAGATGGACTTGAAGGTCCTCTCC	1910
Qy	633	PheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePhe	652
Db	1911	TACTGTCTCAAGCACTGTCTAGAACTTGAAGCAATCCGGGTGGATATCAGAGACCTCTCTC	1970
Qy	653	ProArgAspGluSerAlaGluAlaCysProValValProLeuTrp--MetArgAspLys	671
Db	1971	TCGGTAGAATAACTCTCGAGCTGTGCCCTGTGTGTACTGTCACGAGACACAAATGTAAG	2030
Qy	672	ThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeu	691
Db	2031	CCCTCCTCATGGATGGTGGGAAACTCTGCTCTGCTGTCGCAGCCTCCGGAACCTG	2090
Qy	692	ArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAla	711
Db	2091	AAGGAGCTGCAGTTGGGCGACAGCATCTCGATCAACGGGCCAATGAAGATACTGTGCCTC	2150
Qy	712	LysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThr	731
Db	2151	GAGCTGGGAATCAGTCTCTCGAAATACAGAACTGACGTTTAAAGATGTCAGAGGTAGTG	2210
Qy	732	ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsn	751
Db	2211	TCGTGGCCTGAAACATCTCTGGAAGCTCTCTTTTAGCANTCAAACTTAAAGTACCTCAAT	2270
Qy	752	LeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHis	771
Db	2271	CTAGGGAACACTCCCATGAAGGATGATGATCAATGAAGTTAGCCTCGAAGCGCTGAACAT	2330
Qy	772	ProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyr	791
Db	2331	CCAAAGTGCTCCGTGGAGACTCTGAGGTTGGANTTCCTGTGAGTTAACCATCATTTGGTTAT	2390
Qy	792	LeuLysIleSerGlnIleLeuThrThrSerProSerLysSerLeuSerLeuAlaGly	811
Db	2391	GAGATGATCTCCACGCTCTTATTTCACACACAGCGCTAAAGTGTCTCAGCTGGGCCAAA	2450
Qy	812	AsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCys	831
Db	2451	AATAGTGGGAGTAAAAAGCATGATATCCCTTGGGAATGCCTTGAGTAGCTCAATGTGT	2510
Qy	832	AlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrAlaThrGlyCysGlnSerLeu	851
Db	2511	CTACTGCAAAAGTTGATACTGGCAACTGTGGGCTCACACTCCAGCTGCCACCTCTCTG	2570
Qy	852	AlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeu	871
Db	2571	GTCTCAGCCCTTTTCAGCAACCGAGACTTGACACACTGTGCCCTGTGCNAACAACAGCCTG	2630
Qy	872	GlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGln	891
Db	2631	GGGACTGAAGGAGTGCAACAGCTGTGTGACTTCTTGAGGAATCCAGAATGTGCTCTCCAG	2690
Qy	892	ArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAla	911
Db	2691	CGGCTGATACTGAATCACTCAACATTTGATAGATGCTTATGGCTTCTCTGGCAATGAGA	2750
Qy	912	LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsn	931
Db	2751	CTTGCAAAACACAAAGCTGACCCACTGAGCTGACATGAACCCCGTAGGGATGGT	2810

QY 932 GlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGTACTGTGTGAAGCTTTAAAGAACTACTTGTACCTTCAGAACTGAA 2870
QY 952 LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db 2871 CTAGTGGAGTGCCTCAACTCACAGAACTGCTGCGAGGACCTGGCTGTATGATCAACA 2930
QY 972 SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAAGCACTTAAAGATTTGGATCTTGTGTAACAACGCCCTGGGTGCAAGAGAGTCA 2990
QY 992 AlaLeuCysGluGlyLeuLysGlnLysAsnSerValIleThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGACTGAAGCAAGTAGACAGCTCCCTGAGGAGACTTGGGTGGGGCA 3050
QY 1012 CysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAGTTGACTTCAATGCTGTGAGGCACTTGTCAATGGCCATCTTTCGAACCTCAC 3110
QY 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCys 1051
Db 3111 CTGAACAGCTTAAACCTGTGTAAGAAATGACTTCAGTACATCGGGGATGTTGAAGCTGTC 3170
QY 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuThrTyrGlnTyr 1071
Db 3171 TCTGGTTCATGCTTCAATGCTTCTTAACCTGGGGATAATTTGGCCCTGTGGAAGCAGGATAC 3230
QY 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArgValValIle 1091
Db 3231 TATGCCCGAGTGAGAAGACAGCTGGAGGAAGTTGAGTTTGTCAAGCCCGCTGCTGATT 3290
QY 1092 AspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3291 GATGGTATTGGTAGTGTAGTGTAGTGAAGATGATACCGGAAAC 3329
RESULT 15
AF074018
LOCUS AF074018 3447 bp mRNA linear ROD 10-MAY-2000
DEFINITION Mus musculus MATER protein (Mater) mRNA, complete cds.
ACCESSION AF074018
VERSION AF074018.1 GI:5802697
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Tong, Z. B. and Nelson, L. M.
A mouse gene encoding an oocyte antigen associated with autoimmune
premature ovarian failure
Endocrinology 140 (8), 3720-3726 (1999)
99360614
10433232
PUBMED
2 (bases 1 to 3447)
Tong, Z.-B. and Nelson, L.M.
Direct Submission
Submitted (24-JUN-1998) Developmental Endocrinology Branch, NICHD,
National Institutes of Health, Building 10, Room 10N262, Bethesda,
MD 20892, USA
FEATURES
Location/Qualifiers
1..3447
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
1..3447
/gene="Mater"
9..3344
/gene="Mater"
/note="ooplasm specific protein; OP1"
/codon_start=1
/product="MATER protein"

/protein_id="AAD51762.1"
/db_xref="GI:5802698"
/translation="MGPPKSKAILKARGLEEQSKSRMTSPENDSKSIQKQGPPE
QEGTSTMGPPKSKAILKARGLEEQSKSRMTSPENDSKSIQKQGPPE
ERMTSTMGPPKSKAILKARGLEEQSKSRMTSPENDSKSIQKQGPPE
FDTSDVHSDSPKSKAILKARGLEEQSKSRMTSPENDSKSIQKQGPPE
LFOKSPVLFVREIKWTEKSLAOLIAKECPDSWDVTKINSQPERLLFVLDGLDD
MDSVLDDMTLSRDWKSQPIYILMTSLRLKALLFQSFLLITTRITNLEKLSKMSV
PLYLVEGLSASRRSQVLNENISNDRIOVFHSLIENHOLFQCPQSPVCSLVCBAL
OQKLGKRCCTPCQTLGTYATLVFHLTQVHSLIENHOLFQCPQSPVCSLVCBAL
WTMSVYFDDDLNKLKSESEIALLPHMNLILQVHNSQCVFVSHLSQDPFAALY
VLEGLREMNQHFCEIENORSIMEVKTDDTRLLGMRFLFGLMKNKILKLEVLREVP
VIPTQKQLQHWLIIAQVNGTSPNDTLDFAFYCLFESQDEPFGALKRFPQWILLI
NQMDLKVSYCLHKGCONLKAIRDLLSDVNTLELCPVTQVQCKPLMEVWG
NFCVSLGSLRNKELDGLSDIISQRAKILCLELRNSCRIQKLFKSAFVSVGKHL
WKLLFNQNLKYLNGTMMKDDMKLACEALKHPKCSVETRLDSCETIIGYEMIS
TLIISITRLKCLSLAKNRVGVSMISLGNALSSMCLLOKLILNDCGLTPASCHLVS
ALFSONLTHLCLSNLSLGTGVOQLCOPLRNPECALORLILNHCNIVDAGYFLAMR
LANNTKLTHLSLTMPVPGAMKLLCEALKEPTCYLQLELVDQCLTQCCEDLACMI
TTTHKLSLDLNNALDGVITLCBGLKQSSSLRLGLGKACUKLTSNCCALSLAIS
CNPHLSNLNLVKNDFSTGMLKLCFAFQCPVSNLGIIGLWKQEYARVRRLQLEVEFV
KPHVITDGDWYASDEDDRNWKN"

Alignment Scores: 4.54e-192 Length: 3447
Pred. No.: 2688.00 Matches: 554
Score: 66.49% Conservative: 186
Percent Similarity: 49.78% Mismatches: 293
Best Local Similarity: 38.00% Indels: 81
Query Match: 10 Gaps: 11
DB: 11

US-10-066-521-6 (1-1344) x AF074018 (1-3447)

QY 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPhelLysGluLeuLys 34
Db 163 CAATGGTCTTCCA-GAAAAAGACAGTAAAGCAATCTTGAAGCAGCGTGGATGGAGAG 221
QY 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluLeuLeuAsn 54
Db 222 GAACAGAGTCAAGAAAGCACA----- 242
QY 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
QY 75 SerIleSerIlePheGluAsnMetAsnLeuLeuAspLysGluGluPheGlnThrPhelLysGluLeuLys 94
Db 243 ---ATGTCCTCTTCAAGAAATGTCAGTAGCAGCAATCTTGAAGCAGCGTGGATGGAGAG 299
QY 95 MetLysLysIleSerGln----- 100
Db 300 GTGGAACAGCGGTGCAAGAAAGAAATGACTTCTCCAGAAAAACGACAGTAAATCAATCCAG 359
QY 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 AAGACCAAGGACAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 419
QY 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GAAGTCACAGAGGCA-----GATAAGAT 443
QY 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 444 AATGGAGGTGACTTACAGACTACAGAGGCCATGTGATTGCTTAAGTTCGACACAGAGTG 503
QY 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACGCCACAGAGATGAATTTATTGTCT 542
QY 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 543 GATGCTTTTAAACCATACCAGAAACCTTCCAGCCTCACACCATATCTACATGGAAGA 602


```
Db      2751  CTGCAACACACAAAGCTGACCCACCTGACCTGACCATGAACCCCGTAGGGGATGGT 2810
Qy      932  GlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db      2811  GCAATGAAGCTACTGTGTGAAGCTTTAAAGGAACCTACTTGTACCTTCAAGAACTGGA 2870
Qy      952  LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db      2871  CTAGTGGACTGCCAACTCACACAACTGCTGCAGGACCTGGCTGTATGATCACAACA 2930
Qy      972  SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAla 991
Db      2931  ACCAAGACATTAAAGATTGGATCTTGGTAAACACGCCCTGGGTGACAAAGAGTCTATA 2990
Qy      992  AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db      2991  ACCCTGTGTGAGGGACTGAAGCAAAGTAGCAGCTCCCTGAGGAGACTTGGGTTGGGGCA 3050
Qy      1012  CysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db      3051  TGTAAATTGACTTCCAATTGCTGTGAGGCATGTGCTATGGCCATCTCTGCAACCCCTC 3110
Qy      1032  LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetLysLeuCys 1051
Db      3111  CTGAACAGCCTAAACCTGGTGAAGAAATGACTTCAGTACATCGGGGATGTTGAAGCTGT 3170
Qy      1052  SerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyr 1071
Db      3171  TCTGCGTTCCAATGCCCTGTCTCTAACTGGGGATAATTGGCCTGTGGAGCAGGAGTAC 3230
Qy      1072  ProValGlnIleArgLysLeuLeuGluGluValGlnLeuLysProArgValValIle 1091
Db      3231  TATGCCCGAGTGAGAGACAGCTGGAGGAAGTTGAGTTTGTCAAGCCCCACGTCGTGATT 3290
Qy      1092  AspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db      3291  GATGGTATTGGTATGTAGTGTAGTGAAGATGACCCGAAAC 3329
```

Search completed: July 16, 2004, 04:26:17
Job time : 12554 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 09:07:16 ; Search time 7396 Seconds
(without alignments)
5426.551 Million cell updates/sec

Title: US-10-066-521-6
Perfect score: 1344
Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSNLGAAGLEGVLVS 1344

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n_model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US10066521/runat_13072004_122317_11548/app_query.fasta_1.1543
-DB=EST -QFWT=fastap -SUFFIX=Oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn.1.1.5759 @runat_13072004_122317_11548 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	226	16.8	791	14	CK000513	CK000513 AGENCOURT
2	107	8.0	693	13	BU634350	BU634350 UI-H-FL1-
3	99	7.4	684	13	BU630481	BU630481 UI-H-FL0-
4	85	6.3	642	13	BU618831	BU618831 UI-H-FL1-
5	38	2.8	526	28	AQ542616	AQ542616 RPCI-11-3
6	20	1.5	523	28	AQ428788	AQ428788 CJTBI-E1-
7	20	1.5	814	14	CB228956	CB228956 AGENCOURT
8	16	1.2	743	29	AG117567	AG117567 Pan t598
9	12	0.9	706	10	BB624558	BB624558 BB624558
10	12	0.9	1913	11	AK016782	AK016782 Mus muscu
11	12	0.9	3475	11	AK087774	AK087774 Mus muscu
12	11	0.8	394	14	CF916020	CF916020 B0988G02-
13	11	0.8	511	14	CF173704	CF173704 B0927G04-
14	11	0.8	533	14	CA559680	CA559680 K0261H12-
15	11	0.8	546	14	CA561388	CA561388 K0286C11-
16	11	0.8	559	14	CA561556	CA561556 K0288H05-
17	11	0.8	583	14	CA559979	CA559979 K0266B01-
18	11	0.8	589	14	CA559886	CA559886 K0264H01-
19	11	0.8	593	14	CA558845	CA558845 K0250F09-
20	11	0.8	599	14	CF915731	CF915731 B0984E03-
21	11	0.8	636	14	CF914959	CF914959 B0971H02-
22	11	0.8	643	14	CF913803	CF913803 B0954D05-
23	11	0.8	664	29	AG067278	AG067278 Pan trogl
24	11	0.8	1119	12	BM454123	BM454123 AGENCOURT
25	10	0.7	250	28	BZ170794	BZ170794 CH230-247
26	10	0.7	355	10	BF772019	BF772019 IL5-IT002
27	10	0.7	466	12	BM192717	BM192717 da120H03
28	10	0.7	515	10	BE726276	BE726276 89409D08
29	10	0.7	550	28	BZ302590	BZ302590 K01924.D1
30	10	0.7	754	14	CA817466	CA817466 CAL2E1203
31	9	0.7	172	9	AV421688	AV421688 AV421688
32	9	0.7	192	11	CNS090HH	EX036705 Single re
33	9	0.7	199	12	BJ242372	BJ242372 BJ242372
34	9	0.7	206	10	AW686846	AW686846 NF003B01R
35	9	0.7	236	12	BJ240882	BJ240882 BJ240882
36	9	0.7	238	28	AZ343507	AZ343507 IM0077105
37	9	0.7	263	12	BJ235665	BJ235665 BJ235665
38	9	0.7	264	9	AA231647	AA231647 CD0328.R
39	9	0.7	300	12	BJ262094	BJ262094 BJ262094
40	9	0.7	300	28	B49653	B49653 CIT-HSP-242
41	9	0.7	301	12	BJ308866	BJ308866 BJ308866
42	9	0.7	302	9	AA853292	AA853292 NHTCAe04
43	9	0.7	311	13	BU992299	BU992299 HD09M01R
44	9	0.7	319	12	BJ228111	BJ228111 BJ228111
45	9	0.7	323	10	BB123682	BB123682 BB123682

ALIGNMENTS

RESULT 1
CK000513 791 bp mRNA linear EST 26-NOV-2003
LOCUS AGENCOURT_16368905 NIH_MGC_221 Homo sapiens cDNA clone
DEFINITION IMAGE:30708637 5', mRNA sequence.
CK000513
ACCESSION CK000513.1 GI:38526547
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 791)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-remail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Place: NDA1075 row: c column: 14
 High quality sequence stop: 688.
 Location/Qualifiers

FEATURES

source

```

1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708637"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 221"
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI; Site 2: NotI; Library is oligo-dT primed and directionally cloned. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 4-5Kb. Adaptors 5' (AATTCGGCAGGAGG)3' and 5'd (CTCTGGCCG)3'. 3' linker sequence - GCGGCCGCTGAGAGCC T18. Sequencing primers 3' end: T3 promoter primer 5'd (ATTAACTCCTCAATAAGGGA)3'. 5' End: T7 promoter primer 5'd (TATACAGCTCACTATAGG)3'. Library was constructed in the laboratory of M. Bento Soares. Note: this is a NIH_MGC Library"

```

ORIGIN

Alignment Scores:

Pred. No.:	3,24e-227	Length:	791
Scores:	226.00	Matches:	242
Percent Similarity:	99.59%	Conservative:	0
Best Local Similarity:	99.59%	Mismatches:	1
Query Match:	16.82%	Indels:	1
DB:	14	Gaps:	0

US-10-066-521-6 (1-1344) x CK000513 (1-791)

Qy	190	ValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeu	209
Db	8	GTGGTCTTCGACGGAAGTACGGAATGGGAATCGCTCTAGCCAGAGGATGTCGTG	67
Qy	210	CysTrpAlaGlnGlyLeuTyGlnGlyMetPheSerTyValPheLeuProVal	229
Db	68	TGCTGGCGCAAGGTGGACTTACAGGGAATGTTCTCTACGTCTTCTCTCCCGGTT	127
Qy	230	ArgGluMetGlnArgLysLysSerValThrGluPheLeuSerArgGluTrpPro	249
Db	128	ACAGAGATGCAGCGAAG	187
Qy	250	AspSerGlnAlaProValThrGluLeuMetSerArgProGluArgLeuLeuPheIle	269
Db	188	GACTCCCGAGGTCCTCGGTACCGAGATCATGTGCCACCAAGAGCTGTGTTTCATCAT	247
Qy	270	AspGlyPheAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrp	289
Db	248	GACGGTTTCGATGACCTGGGCTCTGCTCTCAACATGACACAAAGCTCTGCAAGACTGG	307
Qy	290	AlaGluLysGlnProProPheThrLeuLeuArgSerLeuLeuArgLysValLeuLeuPro	309

Db	308	GCTGAGAGAGAGCTCGTTCACCTCATACGAGCTCTGCTGAGGAAGTCTGCTCCCT	367
Qy	310	GluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluVal	329
Db	368	GAGTCTCTCTGATCGTACCGTACAGAGCGTGGGCACAGAGAAGCTCAAGTCAGAGTC	427
Qy	330	ValSerProArgTyrlleuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeu	349
Db	428	GTGTCTCCCGCTACCTGTAGTAGGAATCTCCGGGGAAACAAAGAAATCCACTTGTCTC	487
Qy	350	LeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsn	369
Db	488	CTTGAGCGCGGATGTGTAGCATCAGAGACACAGAGGTGTGTCGATCATGAACAAC	547
Qy	370	ArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLeuLeuValAlaLeu	389
Db	548	CGTGAGCTGCTCGACAGTGCAGGTGCCCGCGTGGGCTCTCTCATCTGCTGGGCGCTG	607
Qy	390	GlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeu	409
Db	608	CAGCTCGAGACGTGTGGGGAGAGCGTGCCTTCAACCAACGCTTCACAGGCTG	667
Qy	410	HisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsn	429
Db	668	CAGCCGCTTTGTGTTAT-CAGCTACCCCTCGAGGCGTGTCCGGCGCTGTCTCAAT	726
Qy	430	LeuGluGlu 432	
Db	727	CTGAGAGAA 735	

RESULT 2
 BU634350/c
 LOCUS
 DEFINITION UI-H-FLI-bgx-m-16-0-UI.s1 NCI CGAP FLI Homo sapiens cDNA clone
 UI-H-FLI-bgx-m-16-0-UI 3', mRNA sequence.
 ACCESSION BU634350
 VERSION BU634350.1 GI:23301605
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 693)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bentso-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

```

1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FLI-bgx-m-16-0-UI"
/tissue_type="Cell lines"
/dev stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FLI"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FLI is a normalized cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was

```

constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGGTG. The cell lines were provided by Dr. James Martin from the University of Iowa.

TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FLI
TAG_SEQ=GAGTCCGGTG

ORIGIN

Alignment Scores:
Pred. No.: 1,586-101 Length: 693
Score: 107.00 Matches: 107
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.96% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x BU634350 (1-693)

Qy 997 LeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSer 1016
Db 642 CTGAAGCAAAAGACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTCTCT 583
Qy 1017 AspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsn 1036
Db 582 GATTGCTGTGAGGCACCTCTCTTGGCCCTTCTCGAACCGGCATCTGACCATCTAAAC 523
Qy 1037 LeuValGlnAsnAsnPheserProLysGlyMetMetLysLeuCysSerAlaPheAlaCys 1056
Db 522 CTGGTGCAGATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCCTTTGGCTGT 463
Qy 1057 ProThrSerAsnLeuGlnLeileGlyLeuThrPlySerGlnTyrProValGlnLea 1076
Db 462 CCCACGCTTAACCTTACAGATAATGGGCTGTGGAAATGCGATACCCCTGTGCAATAAGG 403
Qy 1077 LysLeuLeuGluValGlnLeuLysProArgValValLeuAspGlySerTrpHis 1096
Db 402 AGCTGCTCGAGGAGTGCAGTACTCAAGCCCCGAGTCTGATTTGACGGTAGTTGGCAT 343
Qy 1097 SerPheAspGluAspAspArg 1103
Db 342 TCTTTTGATGAAGATGACCGG 322

RESULT 3

BU630481/c
LOCUS
DEFINITION
UI-H-FLI-bdk-c-22-0-UI-s1 NCI CGAP FLI0 Homo sapiens cDNA clone
UI-H-FLI-bdk-c-22-0-UI 3', mRNA sequence.

ACCESSION
BU630481

VERSION
BU630481.1 GI:23297486

KEYWORDS
EST.

ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 684);
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FLI-bdk-c-22-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DHI0B (Life Technologies)"
/clone_lib="NCI CGAP FLI0"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FLI0 is a cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGGTG. The cell line was provided by Dr. James Martin from University of Iowa."
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FLI
TAG_SEQ=GAGTCCGGTG

ORIGIN

Alignment Scores:
Pred. No.: 4,41e-93 Length: 684
Score: 99.00 Matches: 99
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.37% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x BU630481 (1-684)

Qy 990 ValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeu 1009

Db 656 GTTCTCGCTGTGCGAGGAGCTGACCAAAAGACAGTGTCTGACGAGACTCGGGTTG 597

Qy 1010 LysAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsn 1029

Db 596 AAGGCATGTGAGCTGACTTCTGATTGCTGTGAGGCACCTCTCTGGCCCTTCTCTGCAAC 537

Qy 1030 ArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheserProLysGlyMetMetLys 1049

Db 536 CGGCATCTGACCACTCTAAACCTGGTCAGATAACTTCAGTCCCAAGGAATGATGAAG 477

Qy 1050 LeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnLeileGlyLeuThrLysTrp 1069

Db 476 CTGTGTCGGCCTTGTGCTGTCCACGCTTACACTACAGATAATGGGCTGTGGAAATGG 417

Qy 1070 GlnTyrProValGlnLeileArgLysLeuLeuGluValGlnLeuLysProArg 1088

Db 416 CAGTACCTCTGTGCAAAATAAGGAAGCTCTGTGAGGAGTGCAGCTACTCAAGCCCCGA 360

RESULT 4
BU618831/c

LOCUS
DEFINITION
UI-H-FLI-bfl-i-19-0-UI-s1 NCI CGAP FH1 Homo sapiens cDNA clone

UI-H-FLI-bfl-i-19-0-UI 3', mRNA sequence.

ACCESSION
BU618831

VERSION
BU618831.1 GI:23285046

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 642)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..642
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FH1-bf1-i-19-0-UI"
 /tissue type="Cell Line"
 /dev stages="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_FH1"
 /note="Organ: Chondrosarcoma; Vector: pTT73-Pac
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP_FH1 is a normalized CDNA library
 obtained from a cell line derived from grade I
 chondrosarcoma tissue. The library was constructed and
 normalized according to Bonaldo, Lennon and Soares, Genome
 Research, 6:791-806, 1996. First strand cDNA synthesis was
 primed with an oligo-dT primer containing a Not I site.
 Double stranded cDNA was ligated to an EcoR I adaptor,
 digested with Not I, and cloned directionally into
 pTT73-Pac vector. The oligonucleotide used to prime the
 synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AGAATCCGGC. The cell line was provided by Dr. James Martin
 from the University of Iowa.
 TAG TISSUE=Human Chondrosarcoma Cell Line C8 - Grade 1
 Chondrosarcoma
 TAG_LIB=UI-H-FH1
 TAG_SEQ=AGAATCCGGC"

ORIGIN

Alignment Scores:
 Pred. No.: 2,566-78 Length: 642
 Score: 85.00 Matches: 85
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6,32% Indels: 0
 DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x B0618831 (1-642)

Qy 1004 LeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGluAlaLeuSer 1023
 Db 614 CTGACGACCTCGGGTGAAGGCGATGGACTGCTCTGATTCCTGTGAGGCACTCTCC 555
 Qy 1024 LeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSer 1043
 Db 554 TTGGCCCTTTCTCGCAACCGGCATCTGCACAGTCTAAACCTGGTGCGAGATAACTTCAGT 495
 Qy 1044 ProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIle 1063
 Db 494 CCCAAGAAGATGATGAGCTGTGTGGCCCTTTGCGCTGTCCACGCTAACTTACAGATA 435
 Qy 1064 IleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGluValGln 1083

Db 434 ATTGGGCTGTGGAAATGGCAGTACCTCTGTCAAATAGGAAGTCTGCTGGAGGAGTGCAG 375
 Qy 1084 LeuLeuLysProArg 1088
 Db 374 CTACTCAAGCCCGA 360
 RESULT 5
 AQ542616/c
 LOCUS
 DEFINITION
 RPCI-11-345A9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-345A9,
 genomic survey sequence.
 ACCESSION
 AQ542616
 VERSION
 AQ542616.1 GI:4873072
 KEYWORDS
 GSS.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 526)
 AUTHORS
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
 Venter,J.C.
 TITLE
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 JOURNAL
 Unpublished (1997)
 COMMENT
 Other_GSSs: RPCI-11-345A9.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genet cs (info@resgen.com). BAC end search page:
 http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..526
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7632104"
 /db_xref="taxon:9606"
 /clone="RPCI-11-345A9"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
 RPCI11 Human Male BAC Library"

ORIGIN

Alignment Scores:
 Pred. No.: 9,526-29 Length: 526
 Score: 38.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2,83% Indels: 0
 DB: 28 Gaps: 0

US-10-066-521-6 (1-1344) x AQ542616 (1-526)

Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
 Db 300 GACTGTGGGATACAGCCACGGTTCAGAGTCTGGCCCTCAGCCCTCTGACACACCGG 241
 Qy 860 SerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGlyValAsn 877
 Db 240 AGCTTGACACACCTGTGCTCTATCCACACAGCCTGGGGAACGAAGGTGTAAT 187

```

RESULT 6
AQ428788      523 bp  DNA  linear  GSS 24-MAR-1999
LOCUS
DEFINITION  CITBI-E1-2560117.TF CITBI-E1 Homo sapiens genomic clone 2560117,
              genomic survey sequence.
ACCESSION   AQ428788
VERSION     AQ428788.1  GI:4496554
KEYWORDS
SOURCE      GSS.
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 523)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE       Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
            Map Building
JOURNAL
COMMENT     Other_GSSs: CITBI-E1-2560117.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES
            source
            Location/Qualifiers
                1..523
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="2560117"
                /sex="male"
                /cell_type="sperm"
                /clone_lib="CITBI-E1"
                /note="Vector: p8LOBAC11; Site_1: EcoRI; Site_2: EcoRI;
            Caltech Human BAC Library D"
ORIGIN
Alignment Scores:
Pred. No.:      9, 93e-10      Length:      523
Score:          20.00      Matches:      20
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     1.49%      Indels:      0
DB:              28      Gaps:          0

US-10-066-521-6 (1-1344) x AQ428788 (1-523)
Qy 116 GluIleSerGlnAlaMeGluGlnGlnGluGlyAlaThrAlaGluThrGluGlnGlnGly 135
Db 22 GAATTCACAGCTATGCAACAAGAGTGCCACAGCAGCAGACAGACAGAAACAAGT 81

RESULT 7
CB228956
LOCUS
DEFINITION  AGENCOURT 11501187 NICHD Rh Ovl Macaca mulatta cDNA clone
              IMAGE:6884760 5', mRNA sequence.
ACCESSION   CB228956
VERSION     CB228956.1  GI:28280534
KEYWORDS
SOURCE      EST.
            Macaca mulatta (rhesus monkey)
ORGANISM    Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
            Cercopitheciinae; Macaca.
REFERENCE   1 (bases 1 to 814)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

```

```

TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL
COMMENT     Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
            cDNA Library Preparation: CLONTECH
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLL)
            DNA sequencing by: Agencourt Bioscience Corporation
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LMLL at:
            http://image.llnl.gov
            Plate: LUCM3135 row: b column: 23
            High quality sequence stop: 235.
            Location/Qualifiers
                1..814
                /organism="Macaca mulatta"
                /mol_type="mRNA"
                /db_xref="taxon:9544"
                /clone="IMAGE:6884760"
                /tissue_type="Ovary"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NICHD Rh Ovl"
                /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
            Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.0-4.0 kb. Tissue pooled from
            pre-pubertal, post pubertal sn menopausal monkeys.
            Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN
Alignment Scores:
Pred. No.:      1, 62e-09      Length:      814
Score:          20.00      Matches:      20
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     1.49%      Indels:      0
DB:              14      Gaps:          0

US-10-066-521-6 (1-1344) x CB228956 (1-814)
Qy 743 AlaAsnArgLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGluAspVal 762
Db 78 GCCAACCGTAACCTAAGATCCCTCAATTGGAGGACCCACCTGAAGGAGGAGTGA 137

RESULT 8
AG117567/c
LOCUS
DEFINITION  Pan troglodytes DNA, clone: PTB-125H04.F, genomic survey sequence.
ACCESSION   AG117567
VERSION     AG117567.1  GI:16738086
KEYWORDS
SOURCE      GSS.
ORGANISM    Pan troglodytes (chimpanzee)
            Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            BAC end sequences of Library PTB
            Unpublished
            2 (bases 1 to 743)
REFERENCE   2
AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            Direct Submission
            Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
            1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbases@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
            PRIMERS

```



```

MEDLINE      20499374
PUBMED      11042159
REFERENCE    3
AUTHORS      Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
              Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
              Sumi,N., Ishii,Y., Nakamura,S., Kizama,M., Nishine,T., Harada,A.,
              Yamamoto,R., Matsuoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
              Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M.,
              Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
              Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
              RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
JOURNAL      Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE      20530913
PUBMED      11076861
AUTHORS      4
              The RIKEN Genome Exploration Research Group Phase II Team and the
              FANTOM Consortium.
              Functional annotation of a full-length mouse cDNA collection
JOURNAL      Nature 409, 585-690 (2001)
REFERENCE    5
AUTHORS      The FANTOM Consortium and the RIKEN Genome Exploration Research
              Group Phase I & II Team.
              Analysis of the mouse transcriptome based on functional annotation
              of 60,770 full-length cDNAs
JOURNAL      Nature 420, 563-573 (2002)
REFERENCE    6
AUTHORS      Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
              Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
              Hayashida,K., Hayatsu,N., Hiramoto,K., Hitaoka,T., Hirozane,T.,
              Hori,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
              Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
              Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
              Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
              Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
              Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
              Sogabe,Y., Tagami,M., Tagawa,K., Takahashi,F., Takaki-Akahira,S.,
              Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
              Muramatsu,M. and Hayashizaki,Y.
              Direct Submission
JOURNAL      Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
              RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
              Kanagawa 230-0045, Japan [E-mail:genome-res@sc.riken.go.jp,
              URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
              Fax:81-45-503-9216]
              cDNA library was prepared and sequenced in Mouse Genome
              Encyclopedia Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in RIKEN.
              Division of Experimental Animal Research in Riken contributed to
              prepare mouse tissues.
              Please visit our web site for further details.
              URL:http://genome.gsc.riken.go.jp/
              URL:http://fantom.gsc.riken.go.jp/.
FEATURES      Location/Qualifiers
              source      1..3475
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="FANTOM,DB:E330019F16"
                      /db_xref="WGI:2426482"
                      /db_xref="taxon:10090"
                      /clone="E330019F16"
                      /sex="female"
                      /tissue type="ovary"
                      /clone_lib="RIKEN full-length enriched mouse cDNA library"
                      /dev_stage="2 days pregnant adult"
              misc_feature 77..3366
                      /note="mouse MATER protein
                      (maternal-antigen-that-embryos-require) protein
                      putative"
              polyA_signal 3452..3457

```

```

polyA_site  /note="putative"
            3475
            /notes="putative"

ORIGIN
Alignment Scores:
Pred. No.: 2.3 Length: 3475
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 11 Gaps: 0

US-10-066-521-6 (1-1344) x AK087774 (1-3475)

Qy 861 LeuThrHisLeuCysLeuSerAsnSerLeuGly 872
Db 2620 TTGACACCTGTGCTGTCAAAACACAGCTGGG 2655

RESULT 12
CF916020 394 bp mRNA linear EST 05-NOV-2003
LOCUS B0988G02-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus
DEFINITION musculus cDNA clone NIA:B0988G02 IMAGE:30480745 5', mRNA sequence.
ACCESSION CF916020
VERSION CF916020.1 GI:38187222
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 394)
AUTHORS Plao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from
        submicrogram amounts of total RNAs by a universal PCR amplification
        method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
        Laboratory of Genetics
        National Institute on Aging/National Institutes of Health
        333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
        Email: cdna@lgsun.grc.nia.nih.gov
        Plate: B0988 row: G column: 02
        Seq primer: M13 Reverse
        High quality sequence stop: 394
        POLYA=No.
FEATURES      Location/Qualifiers
              source      1..394
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="niaEST:B0988G02-5"
                      /db_xref="taxon:10090"
                      /clone="NIA:B0988G02 IMAGE:30480745"
                      /dev_stage="Unfertilized Egg"
                      /lab_host="DH10B"
                      /clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long
                      1)"
              note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
              Site 2: NotI; Mouse cDNA project by the Laboratory of
              Genetics, National Institute on Aging (NIA), Intramural
              Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
              This is a long-transcript enriched cDNA library [Ref.
              Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]]. Total
              RNAs were extracted from a pool of 1488 unfertilized eggs.
              Double-stranded cDNAs were synthesized with an Oligo(dT)
              primer [Invitrogen:
              5'-pGACTAGTTCATGATCGGAGCGGCCGCCCTTTTTTTTTTTT-3'],
              treated with T4 DNA polymerase, and purified by
              ethanol-precipitation. The cDNAs were ligated to
              Lone-linker LL-Sal4, purified by phenol/chloroform, and

```

separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 2.35 Length: 394
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CF916020 (1-394)

Qy 500 AlaAlaLeuTyTyValLeuGluGlyLeuGlu 510
Db 140 GTCGCTTATATTATGTTTATAGAGGGCTGGAG 172

RESULT 13

CF173704
LOCUS CF173704 511 bp mRNA linear EST 25-JUL-2003
DEFINITION B0927G04-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus musculus cDNA clone NIA:B0927G04 IMAGE:30474891 5', mRNA sequence.
ACCESSION CF173704
VERSION CF173704.1 GI:33283253
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
1 (bases 1 to 511)
Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0927 row: G column: 04
Seq primer: M13 Reverse
High quality sequence stop: 511
POLYA=No.

FEATURES

source

Location/Qualifiers
1..511
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0927G04-5"
/db_xref="taxon:10090"
/clone="NIA:B0927G04 IMAGE:30474891"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_libs="NIA Mouse Unfertilized Egg cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: Sali; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cdna). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total

RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):
5'-pGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3',
treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 3.14 Length: 511
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CF173704 (1-511)

Qy 500 AlaAlaLeuTyTyValLeuGluGlyLeuGlu 510
Db 271 GTCGCTTATATTATGTTTATAGAGGGCTGGAG 303

RESULT 14

CA559680
LOCUS CA559680 533 bp mRNA linear EST 19-NOV-2002
DEFINITION K0261H12-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus musculus cDNA clone NIA:K0261H12 IMAGE:30050591 5', mRNA sequence.
ACCESSION CA559680
VERSION CA559680.1 GI:25104279
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 533)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Staggs,C.A., Martin,P., Alba,K., Tanaka,T. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library (Long)
Unpublished (2001)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0261 row: H column: 12
Seq primer: M13 Reverse
High quality sequence stop: 533
POLYA=No.

FEATURES

source

Location/Qualifiers
1..533
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:K0261H12-5N"
/db_xref="taxon:10090"
/clone="NIA:K0261H12 IMAGE:30050591"
/tissue_type="Unfertilized Egg"
/lab_host="DH10B"
/clone_libs="NIA Mouse Unfertilized Egg cDNA Library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: Sali; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]}. Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
 5'-pGACTAGTCTAGATCGGAGCGGCCCTTTT-3',
 treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker IL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Alignment Scores:
 Pred. No.: 3.29 Length: 533
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CA559680 (1-533)

Qy 500 AlaAlaLeuTyrValLeuGluGlyLeuGlu 510

Db 427 GCTGCTTATATTATGTTTAGAAGGCTGGAG 459

RESULT 15

CA561388 546 bp mRNA linear EST 19-NOV-2002
 LOCUS K0286C11-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus
 DEFINITION musculus cDNA clone NIA:K0286C11 IMAGE:30052930 5', mRNA sequence.

ACCESSION CA561388

VERSION CA561388.1 GI:25106043

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 546)

REFERENCE Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Staggs, C.A.,

AUTHORS Martin, P., Aiba, K., Tanaka, T. and Ko, M.S.H.

TITLE Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library

(Long)

JOURNAL Unpublished (2001)

COMMENT Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: K0286 row: C column: 11

Seq primer: M13 Reverse

High quality sequence stop: 546

POLYA=No.

FEATURES

source

1. .546 Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="C57BL/6J"

/db_xref="niaest:K0286C11-5N"

/clones="NIA:K0286C11 IMAGE:30052930"

/tissue_type="Unfertilized Egg"

/lab_host="DH10B"

/clone_lib="NIA Mouse Unfertilized Egg cDNA Library

(Long)"

Notes="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001)). [PMID: 11544199]}. Total RNAs were
 extracted from a pool of 1488 unfertilized eggs.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer [Invitrogen]:

5'-pGACTAGTCTAGATCGGAGCGGCCCTTTT-3',
 treated with T4 DNA polymerase, and purified by
 ethanol-precipitation. The cDNAs were ligated to
 Lone-linker IL-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.5 kb. The library was constructed
 by Yulan Piao (NIA)."

ORIGIN

Alignment Scores:
 Pred. No.: 3.38 Length: 546
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CA561388 (1-546)

Qy 500 AlaAlaLeuTyrValLeuGluGlyLeuGlu 510

Db 427 GCTGCTTATATTATGTTTAGAAGGCTGGAG 459

Search completed: July 16, 2004, 15:28:15

Job time : 7413 secs

GentCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 06:26:49 ; Search time 12085 Seconds
(without alignments)
4820.276 Million cell updates/sec

Title: US-10-066-521-6
Perfect score: 1344
Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAGLEGILVS 1344

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q/cn2_1/USPTO_spool/US10066521/runat_13072004_122317_11534/app_query.fasta_1.1543
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521 @CGN 1.1.8359 @runat_13072004_122317_11534 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1:	gb.ba.*	GenEmbl.*
2:	gb.htg.*	
3:	gb.in.*	
4:	gb.om.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sv.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.mu.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	

Result No.	Score	Query Match %	Length	DB ID	Description
1	1006	74.9	3489	6	AX478549 Sequence
2	1006	74.9	3885	9	AY154460 Homo sapi
3	1006	74.9	3926	6	AX704821 Sequence
4	988	73.5	3830	6	AX704823 Sequence
5	968	72.0	5859	6	AX459881 Sequence
6	968	72.0	6939	6	AX459873 Sequence
7	968	72.0	6939	6	AX459891 Sequence
8	616	45.8	3885	9	AY054986 Homo sapi
9	533	39.7	157141	9	AC011470 Homo sapi
10	533	39.7	193609	9	AC024580 Homo sapi
11	385	28.6	1157	6	AX427588 Sequence
12	300	22.3	167509	2	AC012107 Homo sapi
13	300	22.3	167509	2	AC012107 Homo sapi
14	196	14.6	1075	6	AX427590 Sequence
15	169	12.6	167509	2	AC012107 Homo sapi
16	75	5.6	164824	2	AC023887 Homo sapi
17	57	4.2	164824	2	AC023887 Homo sapi
18	55	4.1	1705	9	BC040925 Homo sapi
19	12	0.9	171	10	FI43559S11 Mus muscu
20	12	0.9	3108	6	AX417214 Sequence
21	12	0.9	3221	9	AY116207 Homo sapi
22	12	0.9	3300	6	AX459869 Sequence
23	12	0.9	3303	10	AY329488 Mus muscu
24	12	0.9	3372	10	AY329491 Mus muscu
25	12	0.9	3372	10	AY329485 Mus muscu
26	12	0.9	3375	10	AY329489 Mus muscu
27	12	0.9	3395	9	AY116206 Homo sapi
28	12	0.9	3405	10	AY329487 Mus muscu
29	12	0.9	3432	10	AY329484 Mus muscu
30	12	0.9	3447	6	AX427592 Sequence
31	12	0.9	3447	10	AF074018 Mus muscu
32	12	0.9	3453	10	AY329486 Mus muscu
33	12	0.9	3454	10	AY329490 Mus muscu
34	12	0.9	3466	6	AX833594 Sequence
35	12	0.9	3466	9	AK095460 Homo sapi
36	12	0.9	3470	10	AY196362 Mus muscu
37	12	0.9	3480	10	AY196361 Mus muscu
38	12	0.9	3507	9	BC028069 Homo sapi
39	12	0.9	3534	10	BC053384 Mus muscu
40	12	0.9	3563	9	AY116205 Homo sapi
41	12	0.9	3731	9	AY116204 Homo sapi
42	12	0.9	3827	9	AY095146 Homo sapi
43	12	0.9	3827	9	AY154467 Homo sapi
44	12	0.9	4931	6	AX684291 Sequence
45	12	0.9	67802	2	AC101178 Mus muscu

ALIGNMENTS

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

AX478549
LOCUS AX478549 3489 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 3 from Patent WO248362.
ACCESSION AX478549
VERSION AX478549.1 GI:22217318
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Ramkumar,J. and Arvizu,C.
TITLE Embryogenesis associated proteins
JOURNAL Patent: WO 0248362-A 3 20-JUN-2002;
INCYTE Genomics, Inc. (US)
FEATURES
source
1. 3489
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474830CB1"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3489
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX478549 (1-3489)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 454 ATTTCAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAAGAAATT 513
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 514 TCACAAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAAGACATGGA 573
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 574 GGTGACACATGGGACTACAAAGTCACTGATGATGACCAATTTCGCTGAGGAGGAGATGTA 633
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 634 CGTCGTAGTTTGAACAACACTGCTGCTGACTGGCGGGAATGCAAAACGTTGGCTGGTCT 693
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 694 TTTGATTCAGACCGGTGGGGCTTCGGGCTCCGACCGGTGTTTCGACGGAAAGTCAGGA 753
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 754 ATTGGGAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 813
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 814 CAGGGAATGTTCTCTACGCTTCTCTCTCCCGCTTAGAGAGATGTCAGCGGAAGAAGGAG 873
Qy 238 SerSerValThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 874 AGCAGTGTFCACAGAGTTTCATCTCAGGGAGGTGGCCAGACTCCCAAGGCTCGCGTACCGGAG 933
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleAspGlyPheAspLeuGlySer 277
Db 934 ATCATGTCCCGACCAAGAAGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
Qy 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 994 GTCCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGGAGAGAGCCCTCCGTTCAAC 1053

Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1054 CTATACCGCAGTCTGCTGAGAAAGTCTCTCTCCCTGAGTCTCTCTGATCGTCAACGTC 1113
Qy 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1114 AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTGCTGTCTCCCGTTACCTGTAGTT 1173
Qy 338 ArgGlyLysSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 1174 AGAGGAATCTCCGGGAAACAAAGAATCACTTGTCTCTGAGCGCGGATGTTGGTAGCAT 1233
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgLeuLeuLeuAspGlnCysGln 377
Db 1234 CAGAAGACACAAGGGTTGCTGCGATCATGAACAACCGTGAGCTGCTCGACCAAGTGCAG 1293
Qy 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 1294 GTGCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAG 1353
Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 1354 AGCGTCGCCCTTCAACCAACAGCTCACAGGCTGCACGCCCTTTTGTGTTTCATCAG 1413
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
Db 1414 CTCACCCCTCGAGCGGTGGTCCGGCGTCTCTCAATCTGGAGAAAGAGTGTCTCGAAG 1473
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1474 CGCTTCTCGCGTATGCTGTGGAGGGAGTGTGAAATAGGAAGTCAGTGTTCACGGTGAC 1533
Qy 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 1534 GACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCGTCTCTGTTTCATGACATGAAATC 1593
Qy 478 LeuLeuProAspSerHisCysGluGlyTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db 1594 CTTCTCCAGACAGCACAAGAGTCTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 1653
Qy 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1654 TTTCTGTGCCCTTGTACTACGTGTTAGAGGGCTCGAATTCGAGCAGCTCTCTGCCCC 1713
Qy 518 LeuTyrValGlnLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1714 CTGTACGTTTGAGAAGACAAAGAGTCCATGAGAGTCTTAAACAGGAGGCTTCCATATCCAC 1773
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 1774 TCGCTTTGGATGAAGCGTTTCTTGTGTCCTCGTGAGCGAAGACGTAAGAGGCGCACTG 1833
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 1834 GAGGTCTCTGCTGGGCTCTCCCGTTCCTGAGGGGTGAAGCAGAAAGCTTCTGCACTGGGTC 1893
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 1894 TCTCTGTGGTTCAGCAGCCTAATGCCCACCCAGGAGACACCTCGGACGCTTCCAC 1953
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 1954 TGTCTTTTCAGACTCAAGCAAAAGAGTTGTTGCTTGGCATTTAAACAGCTTCCAGAA 2013
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2014 GTGTGCTTCGATTAACCAAGACCTGGACTTGTATGATGATCTTCTTCTGCTCCAGCAC 2073
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2074 TGTCGCTATTTCGGGAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCC 2133
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677

Db 2134 GCTGAGGCGATGCTCTGCTGCTCTATGATGGATGCGGGATAGACCCCTCATTTGAGGAGCAG 2193
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2194 TGGGAAGATTCTGCTCCATGCTTGGCACCACCCACACCTGCGGCACCTGACCTGGGC 2253
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2254 AGCAGCATCTGACAGAGCGGCCCATGAAGCCCTGTGTGCCAAGCTGAGGCATCCACC 2313
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2314 TGCAAGATACAGACCTCATGTTTGAAGATGCACAGATTACCCCTGGTGTGCAGCACCTC 2373
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 2374 TGGAGATCTGTCATGGCCACACCGTAACTTAAGTCCCTCAACTTGGGAGGCCACCCACTG 2433
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2434 AAGGAAGAGGATGTAAGGATGCGGTGTGAAGCTTTAAACACCCCAAAATGTTTGTGAG 2493
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2494 TCTTTGAGGCTGGATTGCTGTGGATTGACCATGCTCTTACCTGAAGATCTCCCAATC 2553
Qy 798 LeuThrThrSerProSerLeuLysSerLeuLeuAlaGlyAsnLysValThrAspGln 817
Db 2554 CTTAGACCTCCCCAGGCTGAAATCTCTGAGCTGGCAGAAACAAGGTGACAGACAG 2613
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2614 GGAGTAATGCCTCTCAGTGATGCTTGAAGTCTCCAGTGGCCCTGCAGAGCTGATA 2673
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2674 CTGGAGGACTGTGGCATCACAGCCAGCGGTGGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 2733
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 2734 AACCGGAGCTTGACACACTGTGCCCTATCCAAACACGCTGGGGAACGAAGGTGTAAT 2793
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2794 CTACTGTCTGATCCATGAGCTTCCCCACTGTAGTCTGCAGAGGCTGATCTGAATCAG 2853
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2854 TGCCACCTGGACACGCTGGCTGTGGTCTTCTTGCACTTGGCTTATGGGTAACTCATGG 2913
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 2914 CTGAGCACCTGAGCCTTAGCATGAACCTCTGTGAAGACAATGGCGTGAAGCTTCTGTGC 2973
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 2974 GAGTTCATGAGAAACCATCTGTCACTCCAGACCTGGAGTTGGTAAAGTGTCACTC 3033
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3034 ACCGCGCGCTGTGTGAGAGTCTGTCTGTGTGTATCTCGAGGACGACACCTGAAGAGC 3093
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3094 CTGATCTCAGCGCAATGCGCTGGGTGACCGTGGGTGTGCTGACATGTGCGAGGAGCTG 3153
Qy 998 LysGlnLysAsnSerValThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3154 AAGCAAAAGACAGTGTCTTTCAGCAGACTCGGGTTGAAGGCATGTGAGTCACTTCTGAT 3213
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037

Db 3214 TGCTGTGAGGCACTCTCTCTTGGCCCTTTCCTGCACACCGGCATCTGACCGATCTAAACCTG 3273
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3274 GTGCAGATTAACCTTCAGTCCCAAGGAATGATGAGCTGTGTTGCGCTTTCCTGTGCC 3333
Qy 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLys 1077
Db 3334 ACGTCTAATTACAGATAATTTGGCTCTGGAATGGCAGTACCCTGTGCAATAAGGAAG 3393
Qy 1078 LeuLeuGluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSer 1097
Db 3394 CTGCTGAGGAAGTGCAGCTACTCAAGCCCGAGTCGTAATTGACGTTAGTGGCATTTCT 3453
Qy 1098 PheAspGluAspAspArg 1103
Db 3454 TTTGATGAGATGACCCG 3471
RESULT 2
AY154460
LOCUS 3885 bp mRNA linear PRI 20-FEB-2003
DEFINITION Homo sapiens NALP5 (NALP5) mRNA, complete cds.
ACCESSION AY154460
VERSION AY154460.1 GI:28436363
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Tschopp, J., Martinon, F. and Burns, K.
TITLE NALP5: a novel protein family involved in inflammation
JOURNAL Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
MEDLINE 22451042
PubMed 12563287
REFERENCE
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
Location/Qualifiers
1..3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3885
/genes="NALP5"
/genes="NALP5"
1..3603
/notes="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains; similar
to mouse MAYER protein"
/codon_start=1
/product="NALP5"
/protein_id="AA018156.1"
/db_xref="GI:28436364"
/translation="MKVAGGLELGAALISASPRALVTLSTGTCTSLPKNPLPPQNL
SSQPCIMEDKSLTSSYSLQWCLYELDKKEFTFKELKKDKSSSTSCISIQFETE
NANVECLALLHYYGASLAWATSISIFENNNLRTLSSEKARDMMKRHSPPEDPATMGD
QASKEVKGISQAVQDSATAATKEQISQAMEQEGATAETEEQISQAMEQEG
TAAETEEQHGDTWYKSHVMTFAEEDVRRSFENTADWPMTLAGAFSDRWG
FRPRTVLHGKSGIGKSALARRIVLCWAQGLYQGMFSVYFELPVRMQRKKSSTVE
FISREWPDSQAPVTEIMSPERLLFIIDGDDLLGSVINDTLKCKDWAQKQPPFTLIR
SLURKVLPPESFLIVTRDVTBKSKSEVSPRLLVIRGSGISGQRILLERGERHQ
QTGLKRAIMNNRELLDQCQVPASGLICVALQDVGVSAFNPQTLTGLHAAFVHF
QLSPRGVVRCLNLEERVNACFRMAVGVNMRKSVFVGDGDDLMVQGLSESRALFH
MNTLLSDHCEYYTFHLSLODFCAALYYLEGLIEPALCPLYVEKTKRSMELQKA
GFHISLMWKRFLFGLVSDVRPLVLLGCPVLGCVKOKLLHWVSLGQPNATPG
DTLDAFCHLPETQDKSEFVRLALNSFQEWLPIQNQLDLTASSFCLOHCPYLRKRVDP
KGLFPDESAEACFPVPLMWRDKTLLIEQWDFCSMLGTHPLURDLGSSILITERAM
KTCACLRHPTCKIQTMFRNAQITPGVQHLMRIVNMRNLRLNLLGGTHLKEEDVRM
ACBALKPKLLESRLDCCGLTHACYLKIQTLLTSPSLKSLAGNKVTDQGVML
SDALRVSQCALQKLILEDGCTATGCSLASALASVNSRLTHLCLSNNSLNGEWNLLC

MSRLPHCSLORLMNCHLDTPAGCFALALMGNSMLTHLSLWNPVEDNGVKLCE
VWRBPSCHLDLELVKCHLTAAACCESLSCVLSRRHLKSLDLTDLNALODGVAALCSG
LKQNSVLRGLKACAGTSCDCEALSALUSCNRLHLSLNLVQNNFSPKGNMKLCSAF
ACPTSNLQIIGLWKQYEPVQIRKULEEVQLLKPRVWDGSHWSFDEDDRYWKN"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3885
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 9 Gaps: 0

US-10-066-521-6 (1-1344) x AV154460 (1-3885)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGlu 117
Db 568 ATTTCAAGCTATGGAACAAGAGTCCACAGCAGCAGACAGACAGAAAT 627
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGly 137
Db 628 TCACAAGCTATGGAACAAGAGTCCACAGCAGCAGACAGACAGAAATGGA 687
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACATGGAAGTACAGAGTCACTGATGATGACCAATTCGCTGAGGAGGATGTA 747
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 748 CGTCGTAGTTTGAACACACTGCTGCTGACTGCGCGGAAATGCAACGTTGCTGCT 807
Qy 178 PheAspSerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 808 TTTGATTTCAGACCGGTGGGGCTTCGGGCTCGCACGGTGGTTCGACGGAAGTCAGGA 867
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLysTyr 217
Db 868 ATTGGGAATTCGGCTCTAGCAGAGAGTCTGCTGCTGCTGGCGCAAGGTGGACTCTAC 927
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlu 237
Db 928 CAGGGAATGTTCT 987
Qy 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 988 AGCAGTGTTCACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGTAC 1047
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db 1048 ATCATGTCGCCACCGAAGAGCTGTGTTCATCATTTGACGGTTTCGATGACTGGGCTCT 1107
Qy 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 1108 GTCTCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAGAGCAGCCTCCGTTCA 1167
Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1168 CTCATACACAGTCTGCTGAGGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
Qy 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1228 AGAGACGTGGGCACAGAGAACTCAAGTCAGAGGTCTGCTCTCCCGTTACTCTGTTAGT 1287
Qy 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 1288 AGAGGAATCTCCGGGGAACAAGAATCCACTGCTCTTCTGAGCGCGGATTGGTGAGCAT 1347
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 1348 CAGAGACACAGAGGTTCGGTGGCATCATGAAACCGTGCAGCTGCTGCAGACGATGCCAG 1407
Qy 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397

Db 1408 GTGCCCCCGGTGGCTCTCTCATCTGGTGGCCCTCGAGCTGCAGGACGTGGTGGGGAG 1467
Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 1468 AGCGTGGCCCCCTTCAACCAACCGCTCACAGGCTGCACGCCGCTTTGTGTTCATCAG 1527
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
Db 1528 CTCACCCCTCGAGCGGTGGTCCGCGCTGTCTCAATCTGGAGGAAAGAGTGTCTCTGAAG 1587
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1588 CGCTTCGCCGTATGGCTGTGGAGGGAGTGTGAAATAGGAAGTCAGTGTTCAGGGTAC 1647
Qy 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 1648 GACCTCATGTCTCAAGGACTCGGGAGTCTGAGCTCCGTCTCTGTTCACATGAACATC 1707
Qy 478 LeuLeuProAspSerHisCysGluGlyTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db 1708 CTTCTCCACAGACGCCACTGTGTAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 1767
Qy 498 PheCysAlaAlaLeuTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1768 TTTCTGCCCCCTTGTACTACGTGTAGAGGGCTGGAAATCGAGCCAGCTCTCTGCCCT 1827
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1828 CTGTAGCTTCAGAAAGACAAAGAGTTCATGAGAGCTTAAACAGGAGCTTCCATATCCAC 1887
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 1888 TCCTCTTGGATGAAGCGTTTCTTGTGGCTCTGTCGAGCGAAGACGTAAGAGGCCACTG 1947
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTrpVal 577
Db 1948 GAGTCTCTGCTGGGTGTCCTGCTGCTGGGGTGAAGCAGAGCTTCTGCACTGGGTC 2007
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 2008 TCTCTGTGGTTCAGCAGCCTTAATGTCACACCCAGGAGACACCTCGGACGCTTCCAC 2067
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 2068 TGTCTTTTCAGACTCAAGCAAGAGTTTGTGCTTGGCTTGGCATTAACAGCTTCCAGAA 2127
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2128 GTGTGGCTTCGATTAACAGAACCTGGACTTGCATAGCATCTTCTCTGCTCCAGCAC 2187
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2188 TGTCCGTATTTGGGAAAAATTCGGTGGATGTCAAAAGGGATCTTCCCAAGAGATGAGTCC 2247
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
Db 2248 GCTGAGCATGTCTGTGGTCCCTCTATGATGCGGATAAGACCTTCATTGAGAGCAG 2307
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2308 TGGGAAGATTCTCTGCTCATGCTTGGCACCCACCCACACTCGCGGACGCTGGACCTGGC 2367
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2368 AGCAGCATCTGACAGCGCGGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCC 2427
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2428 TGAAGATACAGACCTCTGATGTTTAGAAATGACAGATTACCCCTGGTGTGACGACCTC 2487
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757

```
Db 2488 TGGAGAAATCGTCATGGCCCAACCGTAACCTAAGATCCCTCAACTTTGGGAGCACCACCCACTG 2547
Qy LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2548 AAGGAGAGGATGTAGGATGGCTGTGAACCTTAAACACCCCAAAATGTTGTTGGAG 2607
Qy SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2608 TCTTTAGCGCTGGATGTGCTGGATGACCCATGCCATGCTTACCTGAAGATCTCCCAATC 2667
Qy LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAenLysValThrAspGln 817
Db 2668 CTTAGACCTCCCCAGCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGCCAG 2727
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnSerLeuAlaSerAlaLeuValSer 837
Db 2728 GGAGTAATGCTCTCAGTGATGCTTGGAGTCTCCAGTGGCCCTGCAGAACTGATA 2787
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2788 CTGGAGGACTGTGGCATCACAGCCACGGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 2847
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAenGluGlyValAsn 877
Db 2848 AACCGGAGCTTGACACACTGTGCCTATCCAACAACAGCCTGGGGAACGAAGGTGTAAT 2907
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2908 CTACTGTGTGCATCATGAGGCTTCCCCACCTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2967
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2968 TGCCACCTGGACACGGCTGGCTGTGTTTCTGCACTTGGCTTATGGGTAACTCATGG 3027
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAenGlyValLysLeuLeuCys 937
Db 3028 CTGACGCACCTGAGCCTTAGCATGAACCTCTGGAAGACAATGGCGTGAAGCTTCTGTGC 3087
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluValLysCysHisLeu 957
Db 3088 GAGGTGATGAGAGAACCATCTTGTCATCTCCAGGACCTGGAGTGGTAAAGTGTCATCTC 3147
Qy 958 ThrAlaAlaCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3148 ACCCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGACGACACCTCAAGAGC 3207
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3208 CTGGATCTTCACGGACAATGCCCTGGGTGACGGTGGGGTTGCTGCACCTGTGCGAGGACTG 3267
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3268 AAGCAAAAGAACAGTGTTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGAT 3327
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 3328 TGCTGTGAGGCACTCTCTTGCCCTTCTCTGCAACCGGCATCTGACCACTCTAAACCTG 3387
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3388 GTGCAGAAATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGGCCCTTTGGCCTGCC 3447
Qy 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLys 1077
Db 3448 ACGTCTAACTTACAGATAATTTGGGCTGTGGAATTTGGCAGTAGTACCTGTGCAATTAAGGAG 3507
Qy 1078 LeuLeuGluGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db 3508 CTGCTGGAGGAAGTGCACACTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTTGGCACTCT 3567
Qy 1098 PheAspGluAspArg 1103
Db 3568 TTTGATGAAGATGACCGG 3585
```

```
RESULT 3
AX704821 3926 bp DNA linear PAT 04-APR-2003
LOCUS Sequence 1 from Patent EP1285964.
DEFINITION AX704821
ACCESSION AX704821
VERSION AX704821.1 GI:29561487
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Weiss,B., Lessi,M., Peters-Kottig,M. and Beckmann,G.
TITLE Human mater proteins
JOURNAL Patent: EP 1285964-A 1 26-FEB-2003;
SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES
Location/Qualifiers
source 1..3926
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 6 Gaps: 0
US-10-066-521-6 (1-1344) x AX704821 (1-3926)
Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGluIle 117
Db 454 ATTTCAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAAGAAATT 513
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137
Db 514 TCACAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAAGACATGGA 573
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 574 GGTGACACATGGCAGCTACAGAGTCAGTCATGACCAAAATTCGTCGAGGAGGAGTGA 633
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 634 CGTCGTAGTTTGTAAACACTGCTGCTGACTGGCCGGAATGCAAAACGTTGGCTGCT 693
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 694 TTTGATTACACCGGTGGGGCTTCCGGCCTCGCACGGTGGTCTTCACGGAAGTTCAGGA 753
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 754 ATTTGGGAATCGGCTCTAGCCAGAGGATCGTCTGTGCTGGGGCGCAAGGTGGACTTAC 813
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 814 CAGGGAATGTCTCTACAGTCTTCTTCTCCCGTTAGAGATGTCAGCGGAAGAGGAG 873
Qy 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 874 AGCAGTGTCAAGAGTTCACTCCAGGAGTGGCCAGACTCCCGGCTCGGGTGACGGAG 933
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspLeuGlySer 277
Db 934 ATCATGTCCCGACCAAGAGGCTGTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
Qy 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThr 297
Db 994 GTCTCTCAACATGACACAAAGCTCTGCAAGAAGCTGGGCTGAGAAGCAGCCTCGGTTCA 1053
```

QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB 1054 CTCATACGAGTCTGCTGAGGAGGTCTGCTCCCTGAGTCTCTCTGATCGTACCGTC 1113
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuVal 337
DB 1114 AGAGACGTGGGCACAGAGAAAGCTCAAGTCAGAGGTGCTGCTCTCCCGTTACCTGTTAGTT 1173
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
DB 1174 AGAGAAATCTCCGGGAACAAGAAATCCATGCTCTCTTGGCGGGATTTGGTGAGCAT 1233
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
DB 1234 CAGAAGACACAGAGGTTCGCTGGCATCATGAACAACCGTAGCTGCTCGACAGTGCAG 1293
QY 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
DB 1294 GTGCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGCTGGGGAG 1353
QY 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
DB 1354 AGCGTCGCGCCCTTCAACCAACGCTCACAGGCGCTGCACGCGCTTTTGTCTTCATCAG 1413
QY 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLys 437
DB 1414 CTCACCCCTCGAGGCGTGGTTCGGCGCTGTCTCAATCTGGAGGAAGAGTTGCTCCTGAAG 1473
QY 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
DB 1474 CGCTTCTCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCAGCGGTGAC 1533
QY 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
DB 1534 GACCTCATGGTTCAGAGACTCGGGAGTCTGAGCTCGCTGCTCTGTTTCAATGAACATC 1593
QY 478 LeuLeuProAspSerHisCysGluGluTyrTrpPhePheHisLeuSerLeuGlnAsp 497
DB 1594 CTTCTCCAGACAGCCACTGTGAGGAGTACTACCTTCTTCCACCTCAGTCTCCAGAC 1653
QY 498 PheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
DB 1654 TTCGTGCGCGCTTGTACTACGTGTAGAGGCGCTGGAATCGAGGCGCTCTCTGCGCT 1713
QY 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
DB 1714 CTGTACGTTGAGAACAGAGGTCCATGAGGCTTAAACAGGCGAGGCTTCCATATCCAC 1773
QY 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
DB 1774 TCGCTTTGGATGAGCGTTTCTTTGGCTCTGTGAGCGAAGACGTAAGAGGCCACTG 1833
QY 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
DB 1834 GAGGTCTGCTGGGCTGCTCCCTGCGGGTGAAGCAGAGCTTCTGCACTGGGTC 1893
QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 1894 TCTCTGTTGGGTTCAGACAGCTTAATGCCACCAACCGGAGACACCTCGAGCGCTTCCAC 1953
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 1954 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACAGCTTCCAGAA 2013
QY 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 2014 GTGTGGCTTCCGATTAAACGAACCTTGAGCTTTGATGATCTTCTTCTGCTCCAGAC 2073
QY 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB 2074 TGTCCGTAATTTGCGGAAATTCGGGTGGATGCAAGGGATCTTCCCAAGAGATGAGTCC 2133

QY 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB 2134 GCTGAGGCATGCTCTGTGTCCTCTATGATGCGGGATGAAGCCCTCATTTGAGGAGCAG 2193
QY 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 2194 TGGAGAGATTCTGCTTCATGCTTGGCACCAACCCACACCTGCGGACGTGGACCTGGGC 2253
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 2254 AGCAGCATCTGACAGAGCGGCCATGAACACCTGTGTGCCAGCTGAGGCATCCACC 2313
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 2314 TGAAGATACAGACCTGATGTTTAGAATGACAGATTACCCCTGGTGTGACGACCTC 2373
QY 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
DB 2374 TGGAGAATCGTCAATGCCAACCGTAACTTAAGATCCCTCAACTTGGGAGGCCACCCACTG 2433
QY 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
DB 2434 AAGGAAGAGATGAAGGATGGCGTGTGAAGCCTTAAAAACACCAAAATGTTTGTGGAG 2493
QY 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
DB 2494 TCITTAGAGCTGAGTTGCTGTGATGACCCATGCTTACCTGAGAGATCTCCCAATC 2553
QY 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
DB 2554 CTTACGACCTCCCGGAGCTGAAATCTCTGAGCCTGGCAGAAACAAGGTGACAGACCAG 2613
QY 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
DB 2614 GGAATTAATGCTCTCAGTGTGCTTTCAGAGTCTCCAGTCCGCGCTGCAGAAAGCTGATA 2673
QY 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
DB 2674 CTGGAGACTGTGGCATCACAGCACGGTTCAGAGTCTGGCCTCAGCCCTCGTCAGC 2733
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGGAGCTTGACACACCTGCTTATCCAAACAAGCCTGGGGAACAAGGTGTAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTCATGCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCACCTGGACACGCTGGCTGTGTTTCTTGGCATTTGCGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTGACGACCTGAGCTTGTAGCATGAACCTGTGGAAGACAATGCGCTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeu 957
DB 2974 GAGGTCAATGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
DB 3034 ACCGCGCTGTGCTGAGAGTCTGTCTGTGTGATCTCGAGGAGACACACCTGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCAAGACAAATGCCCTGGGTGACGGTGGGGTGTGTCACCTGTGCGAGGACTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTGAAGGCGATGTGAGCTTCTGAT 3213
QY 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037

```
Db 3214 TCCTGTGAGGCACTCTCTCTTGGCCCTTTCCCTGCAACCGGCATCTGACAGCTTAACACCTG 3273
Qy ValGlnAsnAsnPhSerProlysGlyMetMetLysLeuCySerAlaPheAlaCysPro 1057
Db 3274 GTGCAGATAATCTTCAGTCCCAAGAAATGATGAAGCTGTGTTCGGCCCTTTGCTGTCTCC 3333
Qy 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLys 1077
Db 3334 ACGTCTAACTTACAGATAATTTGGGCTGTGGAATGGCAGTACCCCTGTGCAAAATAAGGAAG 3393
Qy 1078 LeuLeuGluGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db 3394 CTGCTGGAGGAAGTGCAGCTACTCAAGCCCGAGTCGTAATTGACGGTAGTTGGCATCT 3453
Qy 1098 PheAspGluAspAspArg 1103
Db 3454 TTTGATGAAGATGACCGG 3471

RESULT 4
AX704823 3830 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 3 from Patent EP1285964.
ACCESSION AX704823
VERSION AX704823.1 GI:29561488
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Weiss,B., Lessl,M., Peters-Kottig,M. and Beckmann,G.
TITLE Human mater proteins
JOURNAL Patent: EP 1285964-A 3 26-FEB-2003;
SCHERING AKTIENGESSELLSCHAFT (DE)
FEATURES
source
1..3830
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3830
Score: 988.00 Matches: 988
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.51% Indels: 0
Db: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX704823 (1-3830)

Qy 116 GlutSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135
Db 451 GAATTTTCACAGCTATGGAACAGAAAGTCCACAGCAGCAGAGACAGAAACAAGGA 510
Qy 136 HisGlyGlyAspThrTrpAspTrpLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 511 CATGGAGGTGACATGCGGACTACAGAGTACAGTACCAATTCCTGTCAGAGAGAG 570
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 571 GATGTACCTCGTAGTTTGAACACATGCTGCTGACTGGCCGGAATGCAAAACGTTGGCT 630
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 631 GGTGCTTTTGATTCAGACCGGTGGGCTTCGGCTCCGACGGTGGTTCTGCACGGAAG 690
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 691 TCAGGATTTGGGAATCGGCTCTAGCCAGAAAGATCGTGTCTGGCGCGCAAGGTGGA 750
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
```

```
Db 751 CTCCTACAGGGAATGTTCTCTACGTCTTCTTCCCTCCGTTAGAGATGTCAGCGAAG 810
Qy LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 811 AAGGAGAGCAGTGTGACAGAGTTCTATCTCAGGAGGTGGCCAGACTCCCAGGCTCCGGT 870
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeu 275
Db 871 ACGGAGATCATGTCCCGACCCAGAAAGGCTGTGTTCATCATTTGACGGTTTCGATGACCTG 930
Qy 276 GlySerValLeuAsnAsnAspThrLysLeuCyLysAspTrpAlaGluLysGlnProPro 295
Db 931 GGCTCTGTCTCAACATGACACAAAGCTCTGCAAGACTCGGCTGGGTGAGAAAGCAGCTCCG 990
Qy 296 PheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
Db 991 TTACCCCTCATACGAGTCTGCTGAGAAAGGTCCTGCTCCCTGAGTCTTCTCTGATCGTC 1050
Qy 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
Db 1051 ACCGTACAGACAGCTGGGSCACAGAGAGCTCAAGTTCAGAGTCTCGTGTCTCCCGTTACCTG 1110
Qy 336 LeuValArgGlyLysSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGly 355
Db 1111 TTAGTTAGAGGAATCTCCGGGAACAAAGAAATCCACTTGTCTCTTGGAGCGCGGATTTGGT 1170
Qy 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
Db 1171 GAGCATCAGAAGACACAAAGGTTGCGTGCATCATGAACAACCGTGAGCTGCTCGACG 1230
Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
Db 1231 TGCAGGTGCCCCGCGTGGGCTCTCTCATCTGCTGGCCCTGCAGCTGCAGGACGTGGTG 1290
Qy 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
Db 1291 GGGGAGAGCGTCGCCCTTCAACCAACCGCTCACAGCGCTGCACGCGCTTTTGTGTTT 1350
Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValVal 435
Db 1351 CATCAGCTCACCCCTCGAGGCGTGGTCCGCGCTGCTCTCAATCTGGAGGAAAGATTGTC 1410
Qy 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
Db 1411 CTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGATAGGAAGTCAAGTGTTCAC 1470
Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
Db 1471 GGTGACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTTTCACATG 1530
Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeu 495
Db 1531 AACATCTCTCCACAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTC 1590
Qy 496 GlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluLeuProAlaLeu 515
Db 1591 CAGGACTTCTGTGCCCTTGTACTACGTGTAGAGGGCTCGGAATCGAGCCAGCTCTC 1650
Qy 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db 1651 TGCCCTCTGTACGTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCAT 1710
Qy 536 IleHisSerLeuTrpMetLysArgPhePheGlyLeuValSerGluAspValArgArg 555
Db 1711 ATCCACTCGGTTTGGATGAAGCGCTTTCTTGTGTTGGGCTCGTGGAGCGAAGACGTAAGGAGG 1770
Qy 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
Db 1771 CCACTGGAGGTCCTGCTGGGCTGCTCCGTTCCCTGGGGGTGAAGCAAGCTTCTGCAC 1830
Qy 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
```

Db 1831 TGGGTCTCTCTGTTGGGTGACGAGCCTAATGCCACCCAGGAGACACCCCTGAGCC 1890

Qy 596 PheHisCysLeuPheGluThrClnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615

Db 1891 TTCCACTGCTCTTTGAGACTCAAGACAAAGAGTTTGTCTGCTTGGCAATTAAACAGCTTC 1950

Qy 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635

Db 1951 CAAGAGTGTGGCTTCCGATTAAACAGAACTGTGACTTGTATGATCTTCTCTTCTGCTC 2010

Qy 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655

Db 2011 CAGCACTGCTCGGTATTTCCGAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGAT 2070

Qy 656 GluSerAlaGluAlaCysProValValProIleuTrpMetArgAspLysThrLeuIleGlu 675

Db 2071 GAGTCCGCTGAGGCAATGCTCTGCTGCTTATGATGCGGATAAGACCTCATTTGAG 2130

Qy 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695

Db 2131 GAGCAGTGGAGAGATTTCTGCTCAATGCTTGGCACCCACACCACTTGGGCAGCTGGAC 2190

Qy 696 LeuGlySerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLeuArgHis 715

Db 2191 CTGGGCAGCAGCATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCAT 2250

Qy 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735

Db 2251 CCCACCTGCAGATACAGACCTCATGTTTGAATGCACAGATTACCCCTGGTGTGCAG 2310

Qy 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThr 755

Db 2311 CACCTCTGGAGAGATCGTCAATGGCCAAACCGTAACCTTAGATCCCTCAACTTGGGAGGACCC 2370

Qy 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775

Db 2371 CACCTGAAGGAAGAGAGTGTAAAGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTG 2430

Qy 776 LeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795

Db 2431 TTGAGATCTTTGAGGCTGGATTGTGTGATGTGACCATGCCATGCTTACCTGGAAGATCTCC 2490

Qy 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThr 815

Db 2491 CAATCTTACGACCTCCCCAGGCTGAAATCTCTGAGCCTGGCAGGAAACAGGTGACA 2550

Qy 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835

Db 2551 GACCAGGGAGTAAATGCCTCTCAGTGATGCTTGAAGTCTCCAGTGGCCCTGCAGAAG 2610

Qy 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855

Db 2611 CTGATACTGGAGGACTGTGGGATCAGACCCAGGGTTGCCAGAGTCTGGCCCTCAGCCCTC 2670

Qy 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGly 875

Db 2671 GTACGAAACCGAGCTTGACACACTGTGCCCTATCCAAACACACCTTGGGGAAACAGGT 2730

Qy 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895

Db 2731 GTAAATCTACTGTGTGATCCATCAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTG 2790

Qy 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheIleuAlaLeuAlaLeuMetGlyAsn 915

Db 2791 AATCAGTGCCACCTGGACACGGCTGGCTGTGGTTTCTTTCACCTTGGCTTATGGGTAAAC 2850

Qy 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935

Db 2851 TCATGGCTGACGACCTGAGGCTTAGCATGAAACCTCTGTGGAGAACATGGCTGAACTT 2910

Qy 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955

Db 2911 CTGTGCGAGGTGATGAGAGAACCATCTGTGTCATCTCCAGGACCTGGAGTTGGTAAAGTGT 2970

Qy 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975

Db 2971 CATCTCACCCCGCTGCTGTGAGAGTCTGCTCTGTGTGATCTCGAGGAGACACACCTG 3030

Qy 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995

Db 3031 AAGAGCTGTGATCTCAGGCAATGCTGGGTGACCGTGGGGTGTGCTGCACCTGTGGGAG 3090

Qy 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015

Db 3091 GGACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTGAAGGCATGTGACTGACT 3150

Qy 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035

Db 3151 TCTGATGTGCTGAGGCATCTCCTTGGCCCTTCTGCAACCGGCATCTGACGAGCTA 3210

Qy 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055

Db 3211 AACCTGGTGAGATTAATCTCAGTCCCAAGGAATGATGAGCTGTGTCGGCTTTGCC 3270

Qy 1056 CysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075

Db 3271 TGTCCACGCTTAATCTACAGATAATTGGCTGTGAAATGGCAGTACCCTGTGCAATA 3330

Qy 1076 ArgLysLeuLeuGluValGlnLeuLysProArgValValIleAspGlySerTrp 1095

Db 3331 AGGAAGCTGTGAGGAAGTGCAGCTACTCAAGCCCGAGTCGTAATTGCGGTAGTTGG 3390

Qy 1096 HisSerPheAspGluAspArg 1103

Db 3391 CATCTTTTGATGAAGATGACCGG 3414

RESULT 5

AX459881

LOCUS

AX459881

SEQUENCE 24 from Patent WO0240568.

ACCESSION

AX459881.1

GI:21725649

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Tschopp, J. and Martinon, F.

TITLE

Proteins and dna sequences underlying these proteins used for treating inflammations

JOURNAL

Patent: WO 0240568-A 24 23-MAY-2002;

Apotech Research and Development Ltd. (CH)

FEATURES

Location/Qualifiers

1..5859

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="NALP8/Py12---Py8.dna"

ORIGIN

Alignment Scores:

Pred. No.: 0

Score: 968.00

Length: 5859

Matches: 968

Percent Similarity: 100.00%

Conservative: 0

Best Local Similarity: 100.00%

Mismatches: 0

Query Match: 72.02%

Indels: 0

DB: 6

US-10-066-521-6 (1-1344) x AX459881 (1-5859)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117

Db 2389 ATTTCAAGCTATGAAACAAGAGGTGCCACAGCAGCAGACAGAGAACAAGAAATT 2448

Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137

Db	2449	TCACAAGCTATGGAAACAGAGGTCGCCACAGCAGACAGAGAGCAAGAGGACATGGGA	2508
Qy	138	GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspVal	157
Db	2509	GGTGACACATGGGACTACAGAGTCACTGATGATGACCAAAATTCGCTGAGGAGGAGATGTA	2568
Qy	158	ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla	177
Db	2569	CGTCGTAGTTTTGAAAAACACTGCTGCTGACTGCGCGGGAATGCAAAACGTTGCTGGT	2628
Qy	178	PheAspSerAspArgTrpGlyPheArgProArgThrValLeuHisGlyLysSerGly	197
Db	2629	TTTGATTCAGACCGGTGGGCTTCGGCCTCGCACGGTGGTTCGCAACGGAAGTCAGGA	2688
Qy	198	IleGlyLysSerAlaLeuAlaArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr	217
Db	2689	ATTGGGAATCGGCTCTAGCAGCAAGGATCGTGTGTGCTGGGCGCAAGTGGACTCTAC	2748
Qy	218	GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlu	237
Db	2749	CAGGGAATGTTCTCTACGCTCTTCTTCTCCCGTTAGAGAGATGCAGCGGAAGAGAG	2808
Qy	238	SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu	257
Db	2809	AGCAGTGTACAGAGTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGTGACGGAG	2868
Qy	258	IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySer	277
Db	2869	ATCATGTCCGACCAAGAGCTGTTGTTCAITGATGACGTTTCGATGACTGGGCTCT	2928
Qy	278	ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr	297
Db	2929	GTCCCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGGAGAGCAGCCTCCGTTCA	2988
Qy	298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	2989	CTCATACACAGTCTGTCAGAAAGTCTGCTCCCTGAGTCTCTCTGATGTCACCGTC	3048
Qy	318	ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal	337
Db	3049	AGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTCTGTCCTCCCGTTACTCTGTAGTT	3108
Qy	338	ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis	357
Db	3109	AGAGGAATCTCCGGGAACAAAGATCCACTTGCTCTTGAGCGCGGATTTGGTGAGCAT	3168
Qy	358	GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln	377
Db	3169	CAGAAGACACAAAGGTTTCGTCGCATCATGAACAACCGTGAGCTGCTCGACCAAGTGC	3228
Qy	378	ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu	397
Db	3229	GTGCCCGCGTGGGCTCTCATCTGCGTGGCCCTGACAGCTGCAGAGAGTGGTGGGGAG	3288
Qy	398	SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln	417
Db	3289	AGCGTCGCCCCCTTCAACCAACCGCTCACAGCCTGCAGCGCGCTTTGTGTTTCATCAG	3348
Qy	418	LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys	437
Db	3349	CTCACCCCTCAGGCGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTGAAG	3408
Qy	438	ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp	457
Db	3409	CGCTTCTCCGATGCGCTGTGGAGGAGTGGGAATAGGAAGTCAGTGTGTTGACGGTGAC	3468
Qy	458	AspLeuMetValGlnGlyLeuGluSerGluLeuArgAlaLeuPheHisMetAsnIle	477
Db	3469	GACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGCTGCTGTTTCATGACATC	3528
Qy	478	LeuLeuProAspSerHisCysGluGluTyrTrpThrPhePheHisLeuSerLeuGlnAsp	497

Db	3529	CTTCTCCACAGACAGCCACTGTGAGGAGTAGTACTACACCTTCTTCCACCTCTCCAGGAC	3588
Qy	498	PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro	517
Db	3589	TTCTGTGCGGCTTGACTACTAGTGTAGAGGGCTTGAANAATCGAGCCAGCTCTCTGCGCT	3648
Qy	518	LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis	537
Db	3649	CTGTACCTTTCAGGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCATATCCAC	3708
Qy	538	SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu	557
Db	3709	TCGCTTTGGATGAAGCGCTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAAGAGGCCACTG	3768
Qy	558	GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal	577
Db	3769	GAGTCTCTGGGCTGTCCGTTCCCTCGGGGTGAAGCAGAAGCTTCTGCACCTGGGTC	3828
Qy	578	SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis	597
Db	3829	TCTCTGTGGTTCAGCAGCCTAATGCCACCCAGGAGACACCTTGGACGCGCTTCCAC	3888
Qy	598	CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGluGlu	617
Db	3889	TGCTTTTTCAGACTCAAGACAAGAGTGTGTGCTTGGCATTTAAACAGCTTCCCAAGAA	3948
Qy	618	ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis	637
Db	3949	GTGTGGCTTCGATTAACCAAGACCTGGACTTGTATAGCATCTTCTTCTGCTCCAGCAC	4008
Qy	638	CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer	657
Db	4009	TGTCGCTATTTCGGGAAAATTCGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCC	4068
Qy	658	AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln	677
Db	4069	GCTGAGCATGTCTGTGTCTCTATGATGCGGATTAAGACCTTCATTTGAGGAGCAG	4128
Qy	678	TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly	697
Db	4129	TGGGAGATTTCTGCTCCATGCTTGGCACCCACCCACCTCGCGCAGCTGGACCTGGGC	4188
Qy	698	SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr	717
Db	4189	AGCAGCATCTGCAGACGCGGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCAC	4248
Qy	718	CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu	737
Db	4249	TGCAAGATACAGACCCCTGATGTTTAGAAATGCAAGATTACCCCTGGTGTGCACACCTC	4308
Qy	738	TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeu	757
Db	4309	TGAGAGATCTCATGGCCAAACCGTAACCTAAGATCCCTCACTTGGGAGGACCCACCTG	4368
Qy	758	LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu	777
Db	4369	AAGGAAGAGGATGTAGGATGGCGTGTGAAGCTTAAACACCCCAANAATGTTGTTGGAG	4428
Qy	778	SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle	797
Db	4429	TCTTTGAGGCTGGATTTGCTGTGGATTGACCCATGCTTACCTGAAAGATCTCCCAATC	4488
Qy	798	LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln	817
Db	4489	CTTACGACCTTCCCCAGCCTGAAATCTCTGAGCCTGSCAGGAAACAAGGTGACAGACAG	4548
Qy	818	GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle	837
Db	4549	GGAGTAATGCTCTCTAGTGATGCTTGAGAGTCTCCCAAGTGCAGCTGCAGAGCTGATA	4608
Qy	838	LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer	857
Db	4609	CTGGAGGACTGTGGCATCACAGCCACGGGTTCAGAGAGTCTGGCCTCAGCCCTCGTCAGC	4668

Db 4429 CTCACCCCTCAGGCGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTCTGAAG 4488
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 4489 CGCTTCTCCGCTATGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTGACGGTGAC 4548
Qy 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 4549 GACCTCATGGTTCAGGAGCTCGGGAGTCTGAGCTCCGTCTCTGTCTTTCATGAACATC 4608
Qy 478 LeuLeuProAspSerHisCysGluClyTrpThrPhePheHisLeuSerGlnAsp 497
Db 4609 CTTCTCCAGACAGCCACTGTGAGAGTACACCTTCTTCCACCTCAGTCTCCAGGAC 4668
Qy 498 PheCysAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 4669 TTCTGTGCGCTGTACTAGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCCCT 4728
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 4729 CTGTACGTTGAGAAACAGAGGTCCATGAGGCTTAAACAGGCGAGGCTTCCATATCCAC 4788
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 4789 TCGCTTTTGGATGAAGCGTTTCTGTGTGGCTCTGTGAGCGGAGGACGTAAGGAGGCCACTG 4848
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 4849 GAGTCTCTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGCTTCTGCATCTGGGTC 4908
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 4909 TCTCTGTTGGTTCAGCAGCCTAATGCCACCAACCCAGGAGACACCTCGGAGCGCTTCCAC 4968
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 4969 TGTCTTTTCGAGACTCAAGACAAAGAGTTGTGTGCTTGGCTTGAACAGCTTCCAAAGAA 5028
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 5029 GTGTGGCTTCCGATTACCAAGAACTGGACTGTGATAGCATCTTCTCTGTGCTCCAGCAC 5088
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 5089 TGTCCGATTTTGGGMAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAGATGAGTCC 5148
Qy 658 AlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
Db 5149 GCTGAGGATGCTCTGTGTGCTCTATGGATGGGGATAGACCCCTCATTTGAGGAGCAG 5208
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 5209 TGGGAAGATTTCTGCTCATGCTTGGCACCCACCCACACCTCGCGGAGCTGGACCTGGGC 5268
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 5269 AGCAGCATCTGACAGAGCGGCCCATGAGACCTCTGTGTGCCAAGCTGAGGCATCCACC 5328
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 5329 TGCAAGATACAGACCTGTATTTAGAAATGCACAGATTACCCCTGGTGTGTCAGCACCTC 5388
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeu 757
Db 5389 TGGAGATTCGTATGGCCCAACCGTAACCTAAGATCCCTCAACTTGGGAGGCCACCCCTG 5448
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGGAGAGGATGTAGAGTGGCTGTGAACCTTTAAACACCCCAAAATGTTGTGGAG 5508
Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGGATTGCTGTGGATTGACCCCATGCCATGCTGTACCTGAAGATCTCCCAATC 5568

Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTACGACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAGGTGACAGACCAG 5628
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAAATGCTCTCAGTGATGCTTGGAGAGTCTCCAGTGGCGCTGCAAGACTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGAGGACTGTGGCATCACAGCCACGGGTTCGAGAGTCTGGCCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGAGCTTGACACACCTGTGCTTATCCAAACAAGCCTGGGGAACGAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCACCTGGACACGGCTGGCTGTGTGTTTCTTGCACTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGAGCACCCTGAGCCTTAGCATGMAACCTGTGGAGACAACTGGCGTGAAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 5989 GAGTCTATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCACTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 6109 CTGGATCTCAGGACAAATGCCCTGGGTGACGGTGGGGTGTGCTGCATGTGCGAGGAGACTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAAAGAAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGAGCTTCTGTAT 6228
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGSACTCTCTTTGGCCCTTCTCTGCAACCGGCATCTGACCACTCTAAACCTG 6288
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGAGATTAACCTTCAGTCCCAAGGAATGATGAAGCTGTGTCGGCTTTGCTGTGCTCC 6348
Qy 1058 ThrSerAsnLeuGlnIleLeuGly 1065
Db 6349 ACCTCTAATCTTACAGATAATTGGC 6372

RESULT 7

AX459891
LOCUS AX459891 6939 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 34 from Patent WO0240668.

AX459891
ACCESSION AX459891

AX459891.1 GI:21725654

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1

Tschopp, J. and Martinon, F.

Proteins and dna sequences underlying these proteins used for

treating inflammations

Patent: WO 0240668-A 34 23-MAY-2002;

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES Apotech Research and Development Ltd. (CH)

source

Location/Qualifiers

1..6939

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="NALP13/Py17.cdna"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 6939
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX459891 (1-6939)

QY 98 IleSerGlnAlaMetGluGlnGluGlyValAlaThrAlaAlaGluThrGluGlnGluIle 117
DB 3469 ATTTCACAAAGCTATGGAAACAGAAGGTGCCACAGCAGCAGAGACAGAAACCAAGAAATT 3528
QY 118 SerGlnAlaMetGluGlnGluGlyValAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
DB 3529 TCACAAAGCTATGGAAACAGAAGGTGCCACAGCAGCAGAGACAGAAACCAAGACATGGA 3588
QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
DB 3589 GGTGACACATGGGACTACAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGATGTA 3648
QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
DB 3649 CGTCGTAGTTTGAACACACTGCTGCTGACTGCGCGGAAATGCAAAACGTTGGCTGGTGC 3708
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
DB 3709 TTTGATTCAGACCGGTGGGGCTTCGGCCCTCGACGGTGGTTCTGCAGGAAAGTCAAGA 3768
QY 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
DB 3769 ATTTGGAAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGCGCAAGGTGGACTCTAC 3828
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
DB 3829 CAGGAAATGTTCTCCTAGCTTTCTTCTCCCGGTTAGAGAGATGCAGCGGAAGAAGGAG 3888
QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
DB 3889 AGCAGTGTACAGAGATTCTCCAGGGAGTGGCCAGACTCCCAAGGCTCCGGTGACGAG 3948
QY 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
DB 3949 ATCATGTCCCGACACAGAAAGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 4008
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
DB 4009 GTCCCTCAACAATGACAAAGCTCTGAAAGACTCTGGGCTGAGAAAGACGCTCCGTTCAAC 4068
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB 4069 CTCATAGCAGTCTGCTGAGGAAGTCTGTCTCCTGAGTCCCTTCTTCGTATCGTCAACGTC 4128
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
DB 4129 AGACAGCTGGGCAACAGAAAGCTCAAGTCAAGGTGCTGTCTCCCGGTTACCTGTAGATT 4188
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
DB 4189 AGAGGAATCTCCGGGGAAACAAAGAAATCCACTGTCTCTTGGAGCGGGATTTGGTGACAT 4248
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377

DB 4249 CAGAAGACACAAGGGTTGGTGGCATCATGAACAACCGTGAGCTGCTCGACCACTGCAG 4308
QY 378 ValProAlaValGlySerLeuIleCysValAlaAlaLeuGlnLeuGlnAspValValGlyGlu 397
DB 4309 GTCCCGCCCGTGGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGCTGGGGGAG 4368
QY 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
DB 4369 AGCGTGGCCCTTCAACCAACGCTCACAGGCTGCACCCGCTTTTGTTTCATCAG 4428
QY 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
DB 4429 CTCACCCCTCGAGCGTGGTCCGCGCTGTCTCAATCTGGAGGAAAGAGTTGCTCGAAG 4488
QY 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
DB 4489 CGCTTCTGCCGTATGGCTGTGGAGGAGTGTGAATAGGAAGTCAGTGTGTCGGTGAC 4548
QY 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
DB 4549 GACCTCATGGTTCAAGACATCGGGGAGTCTGAGCTCCGTCTCTGTTTCATATGAACATC 4608
QY 478 LeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAsp 497
DB 4609 CTTCTCCAGACAGCACCTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 4668
QY 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
DB 4669 TTTCTGTGCCCTTGTACTACGTGTAGAGGGCTGGAATCGAGCCAGCTCTCTGCCCC 4728
QY 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
DB 4729 CTGTACGTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGACGCTTCCATATCCAC 4788
QY 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
DB 4789 TCGCTTTGGATGAAGCGTTCTTGTGCTCGTGCAGGACACGCTAAGGAGCCACCTG 4848
QY 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisIleTrpVal 577
DB 4849 GAGGTCTGTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAAAGCTTCTGCACTGGGTC 4908
QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 4909 TCTCTGTGGTTCAGCAGCCTAATGCCACCCAGGAGACACCTTGGACGCTTCCAC 4968
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 4969 TGTCCTTTTCGAGACTCAAGACAAAGAGTTGTTGCTTGGCATTTAAACAGCTTCCAGAA 5028
QY 618 ValTrpLeuProIleAsnGlnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 5029 GTGTGGCTTCGATTAACAGAACCTGGACTTGATAGCATCTTCTTCTGCTCCAGCAC 5088
QY 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB 5089 TGTCCGTATTTCGGGAAAATTCGGGTGGATGTCAAAAGGGATCTTCCCAAGAGATGAGTCC 5148
QY 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB 5149 GCTGAGGCATGTCCTGTGGTCCCTCTATGGATGCGGGATAAGACCTCATTTGAGGAGCAG 5208
QY 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 5209 TGGGAAGATTCTGCTCCATGCTGGCACCCACACCTCGCGGACGCTGGACCTGGGC 5268
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 5269 AGCAGCATCTGCAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCAATCCACC 5328
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 5329 TGCAGATACAGACCTGATTTTAGAATGCACAGATTACCCCTGCTGTGTGACGACCTC 5388

```
Qy 738 TtpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 5389 TGGAGAAATCGTCATCGCCACCGTAACCTTAAGATCCCTCACTTGGGAGGACCCACCTG 5448
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGGAAGAGGAGTGAAGGATGCGGTGTGAAGCCTTAAACACCCCAAAATGTTGTTCGAG 5508
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGGATGCTGTGGATGTGACCCATGCTGTACCTGAAGATCTCCCAATC 5568
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTAGGACCTCCCCAGGCTGAAATCTCTGAGCCTGGCAGAAACAGGTGACAGCCAG 5628
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAATGCCTCTCAGTGATGCTGTGAGTCTCCAGTGTGCGCCTGCAGAACTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGGAGGACTGTGGCATCACAGCCACGGGTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGAGCTTGACACACTGTGCTTATCCAAACACAGCCTGGGGACAGAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTCATCATCAGGCTTCCCACTGTAGTCTGTCAGAGGCTGATCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCCACCTGGACACGGCTGGCTGTGTTTCTTTCGACTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGCACTTGAGCTTAGCATGAACCTGTGGAGACAAATGGCGTGAAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeu 957
Db 5989 GAGGTTCATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTATCTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValLysSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyValAspGlyValAlaAlaLeuLysGluGlyLeu 997
Db 6109 CTGGATCTCAGGACAAATGCTTGGGTGACCGTGGGTGTGCTGCACTGTGCGAGGAGCTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAAAGAACAGTGTCTCAGCAGAGCTCGGGTTGAAGGATGTGGACTGACTTCTGAT 6228
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGCACTCTCTCTGCCCCCTTCTGCAACCGGCATCTGACCACTTAACCCCTG 6288
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGCAGATAAATCTCAGTCCCAAGAAATGATGAAGCTGTGTGCGCCTTTGCGCTGTGCC 6348
Qy 1058 ThrSerAsnLeuGlnIleIleGly 1065
Db 6349 ACGTCTTAATCTACATTAATTGGC 6372
```

RESULT 8

AY054986

LOCUS

DEFINITION Homo sapiens maternal-antigen-that-embryos-require protein (MATER)

AY054986

3885 bp mRNA linear PRI 02-APR-2002

mRNA, complete cds.
AY054986
GI:19882272
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3885)
Tong,Z.B., Bondy,C.A., Zhou,J. and Nelson,L.M.
A human homologue of mouse Mater, a maternal effect gene essential
for early embryonic development
Hum. Reprod. 17 (4), 903-911 (2002)
11922687
PUBMED
2 (bases 1 to 3885)
Tong,Z.-B., Bondy,C.A. and Nelson,L.M.
Direct Submission
Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD,
NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA
Location/Qualifiers
1. 3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
1. 3885
/gene="MATER"
1. 3603
/gene="MATER"
/notes="oocyte-specific protein; similar to Mus musculus
Mater encoded by GenBank Accession Numbers AF074018 and
AF143559-AF143573"
/codon_start=1
/product="maternal-antigen-that-embryos-require protein"
/protein_id="AAL15549.1"
/db_xref="GI:19882273"

source

gene
CDS

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3885
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservativeness: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: 9 Gaps: 0

US-10-066-521-6 (1-1344) x AY054986 (1-3885)

Qy 98 IleserGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117

Db 568 ATTTCACAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAACAGAAATT 627

118 SerGlnAlaMetGluGlnGluValaThrAlaAlaGluThrGluGluGlnGlyHisGly 137
Db TCACAGCTATGGACAAAGAGGTGCCACAGCAGAGACAGAGAAACAGGACATGGA 687
138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db GGTGACACATGGGACTACAAGAGTCAGTGATGACCAAAATTCGCTGAGGAGGAGATGTA 747
158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db CGTCTAGTATTTGAAAAACATCGCTGACTGGCGGGAATGCAAACTGGCTGGTCT 807
178 PheAspSerAspArgTrpGlyPheArgProArgThrValLeuHisGlyLysSerGly 197
Db TTGTGATTCAGACCGGTGGGGCTTCGGCCCTCGCAGCGTGTCTGCAACGGAAAGTCAGGA 867
198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db ATTGGGAATCGGCTTAGCCAGAGGATCGTGTGTGCTGGGCGCAAGGTGGACTCTAC 927
218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlyGlu 237
Db CAGGGAATGTTCTCTACGTCTTCTTCTCCCGCTTAGAGAGATGCAGCGGAAAGAGAG 987
238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db AGCAGTGTCAAGAGATTTCATCTCAGGAGGTGGCCAGACTCCCAAGGCTCCGGTGACCGAG 1047
258 IleMetSerArgProGluArgLeuPheIleAspGlyPheAspLeuGlyLys 277
Db ATCATGTCCTCCAGCCAGAAAGGCTGTGTTTCATATTGACGGTTTCGATGACCTGGGCTCT 1107
278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db GTCCCTCAACAATGACAAAGACTCTGCAAGACTGGGCTGAGAAGCAGCCTCCGTTCAAC 1167
298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db CTCATACACAGTCTGCTGAGGAAGTCTGCTCCCTGAGTCCCTTCCTGATCGTCACCGTC 1227
318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db AGAGCGTGGGCA CAGAGAACTCAAGTCAGAGTCTGTGTCTCCCGTTACCTGTTAGTT 1287
338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHis 357
Db AGAGGAATCTCCGGGGAACAAAGAATCCACTTGCTCTTGAGCGCGGGATTGGTGAGCAT 1347
358 GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysGly 377
Db CAGAAGACACAAGGGTTCGCTGCGCATCAT-CAACAACCGTGAGCTGCTGACCAAGTGCCA 1406
377 nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGly 397
Db GGTGCCCGCGTGGGCTCTCATCTGCTGGTGGCCCTGCAGCTGCAGAGCTGGTGGGGGA 1466
397 userValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db GAGCGTCGCCCCCTTCAACCAAAACGCTCACAGGCTGCAGCGCTTTTGC-GTTTCATC 1525
417 InLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db AGCTACCCCCCTCAGAGGGGTGTGTCGGGGCTGTCTCAATCTGGAGGAAGAGTTGTCTGA 1585
437 ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA 457
Db AGCGCTTCTGCGGTATGGCTGTGAGGGAGTGTGGAAATAGGAAGTCAGTGTTGATGGTG 1645
457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db ACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCAACATGAACA 1705

477 leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA 497
Db TCCTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCAGTCTCAGG 1765
497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db ACTTCTGTGCGGCTTGTACTACGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCC 1825
517 roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db CTCTGTACGTTGAGAAACAAAGAGGTCCATGAGCTTAAACAGGAGGAGCTTCCATATCC 1885
537 isSerLeuTrpMetLysArgPheLeuValSerGluAspValArgProL 557
Db ACTCGCTTTGGATGAAGCGTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAAGAGGCGCAC 1945
557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV 577
Db TGGAGTCTCTGTGGGCTGTCCCGTTCCTGGGGGTGAAGCAGAGAGCTTCTGCACCTGGG 2005
577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db TCTCTCTGTGGGTGACGAGCTTAATGCCACCAACCCAGGAGACACCTGGAGCGCTTCC 2065
597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db ACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGTGTGGCATTAACAGAGCTTCCAAG 2125
617 luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db AAGTGTGGCTTCGATTAAACAGAACCTGAGCTTGTATAGCATCTTCTTCTGCTCCAGC 2185
637 isCysProTyrLeuArgLysIleArgValaspValLysGlyIlePheProArgAspGluS 657
Db ACTGTCCGTATTTCGCGGAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGT 2245
657 erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG 677
Db CCCTCAGGAGATGTCTGTGTCTCTATGAGATCGGGATAGACCCCTCATTTGAGGAGC 2305
677 InTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db AGTGGGAAGATTTCTGCTCCATGCTTGGCACCCACCCACCTGCGCGCAGCTGGACCTGG 2365
697 lysSerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT 717
Db GCAGCAGCATCTTGACAGAGCGGCCCATGAAGACCCCTGTGTGCCAAGCTGAGGCATCCCA 2425
717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db CCTGCAAGATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGTGTGTGCAACACC 2485
737 euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db TCTGGAGATTCGTCATGGCCACCGTAACCTAAGATCCCTCAACTTGGGAGGACCACCCACC 2545
757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db TGAAGGAAGAGGATGAAGAGTGGCGTGTGAAGCCTTAAACACACCCCAAAATTTGTTGTG 2605
777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db AGTCTTTGAGGCTGGGATGCTGTGGATGACCCATGCCATGCCCTGTATTACCTGGAAGATCTCCAAA 2665
797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db TCCTTAGACCTTCCCCAGCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACC 2725
817 InGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db AGGGAGTAAC-GCCTCTCAGTGTATGCTTTGAGGGTCTCCAGTGCCTCCCTGCAAGAGCTG 2784
837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856

Db	2785	ATACTGAGGACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGCCTCAGCCCTCGTC	2844
Qy	857	SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal	876
Db	2845	AGCAACCGGAGCTTGACACACCTGTGCTATCCAAACAACGCTGGGGACGAGGTGTA	2904
Qy	877	AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn	896
Db	2905	AATCTACTGTGTTCATCCATCAGAGCTTCCCACTGTAGTCTGCAGAGGCTCATCTGAA	2964
Qy	897	GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe	916
Db	2965	CAGTGCCACCTGGACACGGCTGGCTGGG-TTCTTTGCACTTGGCTTATGGGTAACTC	3023
Qy	916	rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe	936
Db	3024	ATGGCTGACGCACTTGAGCCTTAGCATGAACCTGTGGAAGACATGGCGTGAAGCTTCT	3083
Qy	936	uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi	956
Db	3084	GTGCGAGGTATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTCTCA	3143
Qy	956	sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy	976
Db	3144	TCTCACCAGCCGCGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAA	3203
Qy	976	sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGl	996
Db	3204	GAGCTGGATCTCACGGACATCGCCCTGGGTGACGGTGGGGTCTGCCGTGTCCGAGG	3263
Qy	996	yLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe	1016
Db	3264	ACTGAAGCAAAAGAACAGTGTCTTGACGAGACTCGGGTTGAAGGCATGTGCACTGACTTC	3323
Qy	1016	rAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs	1036
Db	3324	TGATTTGTGTGAGGACATCTCTTTGGGCCCTTTCCTGCAACGGCATCTGACCACTCTAA	3383
Qy	1036	nLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCy	1056
Db	3384	CTGGTGCAAGATTAACCTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGCC	3443
Qy	1056	sProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpValGlnIleAr	1076
Db	3444	TCCACAGCTCTAACTTACAGATAATTGGGCTGTGGAATGCGAGTACCTCTGCAATAG	3503
Qy	1076	gLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHi	1096
Db	3504	GAAGCTCTCGAGGAAGTGCAGCTACTCAAGCCCGAGTCTTAATTCACGGTAGTTGGCA	3563
Qy	1096	sSerPheAspGluAspAspArg 1103	
Db	3564	TTCTTTTGATGAAGATGACCGG 3585	
RESULT	9		
AX427610			
LOCUS	AX427610	3900 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 23 from Patent WO0232955.		
ACCESSION	AX427610		
VERSION	AX427610.1	GI:21537730	
KEYWORDS	homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Nelson, L.M. and Tong, Z.-B.		
TITLE	Human gene critical to fertility		
JOURNAL	Patent: WO 0232955-A 23 25-APR-2002;		
GOVERNMENT OF THE UNITED STATES (US)	Location/Qualifiers		
FEATURES	1..3900		

Qy	258	IleMetSerArgProGluAtrLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer	277
Db	1048	ATCATGTCCCGACACAGAAAGGCTGTGTTCATCATATTGACGGTTTCGATGACCTGGGCTCT	1107
Qy	278	ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr	297
Db	1108	GTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAGAGCAGCCTCCGTTCAAC	1167
Qy	298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	1168	CTCATACGACTCTGCTGAGGAAGTCTGCTCCCTGAGTCTCTTCCTGATCGTCAACGCTC	1227
Qy	318	ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal	337
Db	1228	AGAGACGTGGGCACAGAGAGCTCAAGTCAGAGTCTGTCTCCCGTTACCTGTTAGTT	1287
Qy	338	ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis	357
Db	1288	AGAGGAATCTCCGGGGAACAAGAATCCACTTGTCTCTTGAGCGCGGATTTGTTGACAT	1347
Qy	358	GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuLeuAspGlnCysGl	377
Db	1348	CAGAAGACACAAAGGTTGCGTGCATCAT-CAACAACCGTGAGTGTCTGCACCAAGTGCCA	1406
Qy	377	nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGl	397
Db	1407	GGTGCCCGCCGTGGGCTCTCTCATCTCGTGCGCCCTGCAGCTGCAGACGTGTGGGGGA	1466
Qy	397	uSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG	417
Db	1467	GAGCGTCGCCCCCTTCAACCAACCGCTCACAGGCTGCACGCCCTTTTGC-GTTTCATC	1525
Qy	417	lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL	437
Db	1526	AGCTCACCCCTCGAGGGGTGTTCGGCGCTCTCTCAATCTGGAGGAAGAGTTGTCTCTGA	1585
Qy	437	ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA	457
Db	1586	AGCGCTCTCGCGTATGCTGTGAGGAGGTGTGGAAATAGGAAGTCAAGTGTTCATGGTG	1645
Qy	457	spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI	477
Db	1646	ACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGAACA	1705
Qy	477	leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisSerLeuGlnA	497
Db	1706	TCCTTCTCCACAGACGCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGG	1765
Qy	497	spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP	517
Db	1766	ACTTCTGTGCGCCCTTGTACTAGTGTAGAGGCTTGGAAATCGAGCCAGCTCTCTGCC	1825
Qy	517	roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH	537
Db	1826	CTCTGTACTGTTGAGAAGACAAGAGGTCCATGGAGCTTAAACAGCGAGGCTTCCATATCC	1885
Qy	537	isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL	557
Db	1886	ACTCGCTTTGGATGAAGCGTTTCTTGTGTGCGCTCGTGAGCGAAGAGCTAAGGAGGCCAC	1945
Qy	557	euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV	577
Db	1946	TGGAGGCTCTCTGGCTGTCTCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGCACGTGGG	2005
Qy	577	alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH	597
Db	2006	TCTCTCTGTTGGGTCAGCAGCTTAATGCCACACCCCGAGGAGACACCTCGAGCGCTTCC	2065
Qy	597	isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG	617
Db	2066	ACTGTCTTTTCGAGACTCAAGACAAAAGATTGTTCGCTTGGCATTAACACAGCTTCCAAG	2125

Qy	617	luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH	637
Db	2126	AAGTGTGGCTTCGATTTAACCAAGACTGAGACTTGATAGATCTTCTTCTTGCTTCAGC	2185
Qy	637	isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS	657
Db	2186	ACTGTCCGATTATTCGGGAAAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGAGATGAGT	2245
Qy	657	erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG	677
Db	2246	CCGCTGAGGCATGCTCGTGTCTCTATGGATGCGGGATAAGACCTCATTTGAGAGC	2305
Qy	677	lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG	697
Db	2306	AGTGGGAGATTTCTGCTCCATGCTTTGGCCACCACCCACACCTCGCGCAGCTGGACTGG	2365
Qy	697	lySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT	717
Db	2366	GCAGCAGCATCTGCACAGAGCGGCCCATGAAGACCTGTGTGTGCCAAGCTGAGGCATCCCA	2425
Qy	717	hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL	737
Db	2426	CCTGCAAGATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGTGTGCACACACC	2485
Qy	737	euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL	757
Db	2486	TCGTGGAGATCGTCATGGCCACCGTAACCTAGATCCCTCAACTTTGGGAGGCACCCACC	2545
Qy	757	eulysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG	777
Db	2546	TGAAGGAAGAGGATGTAAGGATGCGGTGAAGCCTTAAAAACACCCAAAATGTTGTGTGG	2605
Qy	777	luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI	797
Db	2606	AGTCTTTGAGGCTGGAATGCTGTGGAATGACCATGCCTGTTACCTGAAGATCTCCAAA	2665
Qy	797	leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG	817
Db	2666	TCCTTAGCACCTCCCGAGCCTGAANAATCTGAGCCTGGCAGGAAACAAAGGTGACAGACC	2725
Qy	817	lnGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu	836
Db	2726	AGGGAGTAAAC-GCCTCTCAGTGATGCTTGAGGCTCTCCAGTGCCTGCGCAAGAGCTG	2784
Qy	837	IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal	856
Db	2785	ATACTGCAGGACTGTGGCATCACAGCACGGTGTGCCAGAGTCTGGCCTCAGCCCTCGTC	2844
Qy	857	SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGlyVal	876
Db	2845	AGCAACGGGAGCTTGACACACCTGTGCCTATCCAACAACAGCCTGGGGAACGAAGGTGTA	2904
Qy	877	AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn	896
Db	2905	AATCTACTGTGTGCATCCATAGGCTTCCCACTGAGTCTGCAGAGGCTGATGCTGAAT	2964
Qy	897	GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyVense	916
Db	2965	CAGTGGCACCTGGACACGGCTGGCTGTGG-TTCTCTTGACACTTGGCCTATGGGTAACTC	3023
Qy	916	rTrpLeuThrHisLeuSerSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe	936
Db	3024	ATGCTTGACGCACCTGAGCCTTAGCATGAACCTGTGGAGACAAATGGCGTGAAGCTTCT	3083
Qy	936	uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi	956
Db	3084	GTGCGAGGTCAATGAGAGAACCATTCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTC	3143
Qy	956	sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLys	976
Db	3144	TCTACCGCGCGCTGTGTGAGAGTCTGCTCTGTGTGATCTCGAGGAGCAGACACCTGAA	3203
Qy	976	sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGl	996

[illegible]

```
QY 475 MetAenlleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSer 494
Db 40980 ATGACATCCCTTCCTCCGAGACAGCCACTGTGAGGAGTACTACACTTCTCCACCTCAGT 40921
QY 495 LeuGlnAspPheCysAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAla 514
Db 40920 CTCACGAGACTTCGTGGCCGCTTGTACTACCTGTGTAGAGGCGCTGGAAATCGAGCCAGCT 40861
QY 515 LeuCysProLeuTyrValGluIleThrIlySerMetGluLeuLysGlnAlaGlyPhe 534
Db 40860 CTCGCCCTCTGTAGTGTGAGAGCAAAAGAGGTCATGGAGCTTAAACAGGCAGGCTTC 40801
QY 535 HisIleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
Db 40800 CATATCCACTGCCTTGTGATGAGCGTTCCTGTGTGGCTTCGTGAGGAGAGCTAAGG 40741
QY 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
Db 40740 AGGCCACTGGAGGTCCTGCTGGGCTGTCCGCTTCCCTGGGGGTGAAGCAGAGACTTCTG 40681
QY 575 HisTrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAsp 594
Db 40680 CACTGGGTCTCTCTGTGGGTGAGCAGCCTAATGCCACACCCAGGAGACACCCCTGGAC 40621
QY 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
Db 40620 GCCTTCCACTGTCTTTCGAGACTCAAGACAAGAGTTGTTCCTGTGGCATTAACAGC 40561
QY 615 PheGlnGluValTrpLeuProIleAsnGlnLeuAspLeuIleAlaSerSerPheCys 634
Db 40560 TTCCAAGAAGTGTGGCTCCGATTAATACCAGAACCTGGACTTGTAGCATCTTCTCTGCG 40501
QY 635 LeuGlnHisCysProTyrLeuArgLysIleArgValLysGlyIlePheProArg 654
Db 40500 CTCAGACACTGCTCGCTATTTTCGGAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGA 40441
QY 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667
Db 40440 GATGAGTCCGTGAGGCATGCTGTGTGCTCTTAAGG 40402

RESULT 11
AC024580 193609 bp DNA linear PRI 21-DEC-2001
LOCUS Homo sapiens chromosome 19 clone CTD-2621117, complete sequence.
DEFINITION AC024580
ACCESSION AC024580
VERSION AC024580.6 GI:17975240
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Unpublished
AUTHORS 2 (bases 1 to 193609)
JOURNAL DOE Joint Genome Institute.
TITLE Direct Submission
AUTHORS Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Submitted (13-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS Drive, Walnut Creek, CA 94598, USA
COMMENT On Dec 21, 2001 this sequence version replaced gi:16905144.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
```

Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.9.

FEATURES

source
1. 193609
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2621117"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 193609
Score: 533.00 Matches: 533
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.66% Indels: 0
DB: 9 Gaps: 0

US-10-066-521-6 (1-1344) x AC024580 (1-193609)

QY 135 GlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 154
Db 22258 GGCATGGAGGTGACATGGGACTACAGAGTCACGTGATGACCAAAATTCGCTGAGGAG 22317
QY 155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 22318 GAGGATGTACGTGCTAGTGTGAAACACACTGCTGCTGACTGGCCGGAATGCAACGTTG 22377
QY 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
Db 22378 GCTGGTGCTTTTCATTCAGACCGGTGGGCTTCGCGCTCGCACGGTGGTTCGTCACGA 22437
QY 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 214
Db 22438 AAGTCAGGAATTCGGAAATCGGCTCTAGCCAGAGAGATCGTCTGTGCTGGGCGCAAGT 22497
QY 215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
Db 22498 GGACTCTACAGGGAATGCTCTACGTCTTCTCTCCCGTTAGAGAGATGCAGCGG 22557
QY 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
Db 22558 AAGAAGGAGCAGCTGTCACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCG 22617
QY 255 ValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAsp 274
Db 22618 GTGACCGAGATCATGTCCGACCCAGAAAGGCTGTGTTTCATCATTCACGGTTTCGATGAC 22677
QY 275 LeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 22678 CTGGGCTCTGCTCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAAGCAGCCT 22737
QY 295 ProPheThrIleLeuArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 22738 CCGTTACCCCTCATACGAGTCTGCTGAGAAAGGCTCTGCTCCCTGAGTCTTCTCTGATC 22797
QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
Db 22798 GTACCCGTGAGAGCTGGGACAGAGAAAGCTCAAGTCAGAGGTCTGTCTCCCCGTTAC 22857
QY 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 22858 CTGTTAGTTAGAGAAATCTCCGGGAAACAAAGAAATCCACTTGTCTCTTTCGAGCGGGATT 22917
QY 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
Db 22918 GGTGAGCATCAGAAGACACAAGGGTTGCGTGGCATCATGAACAAACCGTAGCTGCTCGAC 22977
QY 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394

Db	22978	CAGTCCAGGTCGCCCGCGTGGGCTCTCTCATCTCGGTGGCCCTGCAGCTCGAGACGTG	23037
Qy	395	ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal	414
Db	23038	GTGGGGGAGAGGTGCGCCCTTCAACCAAGCTCAGAGCTCGACGCGCTTTGTG	23097
Qy	415	PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgVal	434
Db	23098	TTTCATCAGCTCACCCCTCGAGCGGTGGTCCGGCGCTGTCTCAATCTCGAGAAAGATT	23157
Qy	435	ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe	454
Db	23158	GTCTTAAGAGCGTCTTCGCGATAGCTGTGGAGGAGTGTGAAATAGAAATGAGTGT	23217
Qy	455	AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis	474
Db	23218	GACGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTCCTCTTCAC	23277
Qy	475	MetAsnLeuLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSer	494
Db	23278	ATGAACATCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGT	23337
Qy	495	LeuGlnAspPheCysAlaAlaLeuTyrThrValLeuGluGlyLeuGluGluProAla	514
Db	23338	CTCCAGGACTTCTGTGCCGCTTGTACTACGTGTAGAGGGCCCTGGAAATCGAGCCAGCT	23397
Qy	515	LeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe	534
Db	23398	CTCTGCCCTCTGTAGTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGGAGGCTTC	23457
Qy	535	HisLeHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArg	554
Db	23458	CATATCCACTCGCTTTGGATGAAGCGTTCTTGTGTGGCTCGTGAGCGAAGCGTAAGG	23517
Qy	555	ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu	574
Db	23518	AGCCACCTGGAGGTCCTCTCGGCTGTCCCGTCCCTGGGGGTGAAGCAGAGCTTCTG	23577
Qy	575	HisTrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAsp	594
Db	23578	CACCTGGGTCTCTGTGGGTGAGCAGGCTTAATGCCACCACCCAGGAGACCCCTGGAC	23637
Qy	595	AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer	614
Db	23638	GCTTCCACTCTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGTTGGCATTAAACAGC	23697
Qy	615	PheGlnGluValTrpLeuProLeuAsnGlnAsnLeuAspLeuLeuAlaSerSerPheCys	634
Db	23698	TTCCAAGAGGTGGCTTCCGATTAACAGAACCTGGGACTTGATAGCATCTTCTCTCTGC	23757
Qy	635	LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg	654
Db	23758	CTCAGACACTGCTCGTATTTGGGAAATTCGGGTGGATGTCAAGAGGATCTTCCCAAGA	23817
Qy	655	AspGluSerAlaGluAlaCysProValValProLeuTrp	667
Db	23818	GATGAGTCCGCTGAGGATGCTCTGTGGTCCCTCTATGG	23856
RESULT 12			
LOCUS	AX427588	1157 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 1 from Patent WO0232955.		
ACCESSION	AX427588		
VERSION	AX427588.1	GI:21537711	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Nelson, L.M. and Tong, Z.-B.		
TITLE	Human gene critical to fertility		
JOURNAL	Patent: WO 0232955-A 1 25-APR-2002;		

FEATURES		GOVERNMENT OF THE UNITED STATES (US)	
source	1. 1157	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:		Length:	1157
Pred. No.:	0	Matches:	385
Score:	385.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	28.65%	Gaps:	0
DB:	6		
US-10-066-521-6 (1-1344) x AX427588 (1-1157)			
Qy	252	GlnAlaProValThrGluIleMetSerArgProGluArgLeuPheIleLeuAspGly	271
Db	1	CAAGCTCCGGTGACGGAGATCATGTCCTCCGACCAAGAGGCTGTTGTTTCATCATTTGACGGT	60
Qy	272	PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu	291
Db	61	TTGCGATGACCTGGGCTCTGTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG	120
Qy	292	LysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer	311
Db	121	AAGCAGCTCCGTTTACCCCTCATACGAGTCTGCTGAGGAAGGTCTGCTCCCTCCCTGAGTCC	180
Qy	312	PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer	331
Db	181	TTCTGATGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTCTGTCT	240
Qy	332	ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu	351
Db	241	CCCCGTTTACCTGTTAGTTAGGAATCTCCGGGGAAACAAAGAATCCACTTGTCTCTTGAG	300
Qy	352	ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu	371
Db	301	CGCGGAGATTGGTGGAGCATCAGAAGACACAAGGGTGGCTCGGATCATGAACACCGTGAG	360
Qy	372	LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu	391
Db	361	CTGCTCGACAGTCCAGCGTGGCGCGTGGGCTCTCTCATCTCGGTGGCCCTGCAGCTG	420
Qy	392	GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla	411
Db	421	CAGGACGTGGTGGGGAGAGCGTCCGCCCTTCAACCAACCGCTCACAGGCTGCACGCC	480
Qy	412	AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu	431
Db	481	GCCTTTTGTGTTTTCATCAGCTCACCCCTCGAGGGGTGGTCCGGCGCTGTCTCAATCTGGAG	540
Qy	432	GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys	451
Db	541	GAAGAGTTGCTCTTGAAGCGCTTCTCCCGTATGGCTGTGGAGGGAGTGTGGAATAGGAAG	600
Qy	452	SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla	471
Db	601	TCAGTGTTCGAGCGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCT	660
Qy	472	LeuPheHisMetAsnLeuLeuProAspSerHisCysGluGluTyrThrPhePhe	491
Db	661	CTGTTTCATCATGAACATCTCTTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTC	720
Qy	492	HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIle	511
Db	721	CACCTCAGTCTCCAGGACTTCTGTGCGCGCTTGTACTACTGTGTAGAGGGCCCTGGAATC	780
Qy	512	GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln	531
Db	781	GAGCCAGCTCTCTGCTCTGTACGTTGAGAGACAAAGAGGTCCATGGAGCTTAAACAG	840

Qy 532 AlaGlyPheHisIleHisSerLeuTrpMetIysArgPheLeuPheGlyLeuValSerGlu 551
 Db 841 CGAGGCTTCCATATCCACTCGCTTGGATGAAGGTTCTTGTGTGGCCCTGGAGCGAA 900
 Qy 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValIysGln 571
 Db 901 GACGTAGAGGCCACTGGAGTCTCTGCTGGCTGTCCGTTCCCTTGGGGGTGAACAG 960
 Qy 572 LysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrProGlyAsp 591
 Db 961 AAGCTTCTGCACCTGGGTCTCTCTGTGGTCAGCAGCTAATGACCAACCCAGGAGAC 1020
 Qy 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspIysGluPheValArgLeuAla 611
 Db 1021 ACCCTGGAGCGCTTCCACTGCTCTTTCGAGACTCAAGCAAGAAGTTGTTCGCTTGGCA 1080
 Qy 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
 Db 1081 TTAACAGCTTCCAGAGAGTGGCTCCGATTACCAAGACCTGGACTGTATAGCATCT 1140
 Qy 632 SerPheCysLeuGln 636
 Db 1141 TCCTTCTGCTCCAG 1155

RESULT 13
 AC012107/c
 LOCUS Homo sapiens clone RP11-45K21, WORKING DRAFT SEQUENCE, 25 unordered
 DEFINITION pieces.
 AC012107
 AC012107.2 GI:7329252
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 167509)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens, clone RP11-45K21
 Unpublished
 2 (bases 1 to 167509)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouhgalter,B.,
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
 Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Headford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,E.,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Teafaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 26, 2000 this sequence version replaced gi:6088020.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 ftp://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WITR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L1180
 Center clone name: 45_K_21
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads
 Chemistry: Dye-terminator Big Dye; 96% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 145437 bases at least Q40
 Consensus quality: 155496 bases at least Q30
 Consensus quality: 159832 bases at least Q20
 Insert size: 176000; agarose-fp
 Insert size: 165109; sum-of-contents
 Quality coverage: 3.8 in Q20 bases; agarose-fp
 Quality coverage: 4.1 in Q20 base.
 NOTE: This is a 'working draft' sequence. It currently
 * consists of 25 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1061: contig of 1061 bp in length
 * 1062 1161: gap of 100 bp
 * 1162 2381: contig of 1220 bp in length
 * 2382 2481: gap of 100 bp
 * 2482 3771: contig of 1290 bp in length
 * 3772 3871: gap of 100 bp
 * 3872 4949: contig of 1078 bp in length
 * 4950 5049: gap of 100 bp
 * 5050 6119: contig of 1070 bp in length
 * 6120 6219: gap of 100 bp
 * 6220 8023: contig of 1804 bp in length
 * 8024 8123: gap of 100 bp
 * 8124 10555: contig of 2432 bp in length
 * 10556 10655: gap of 100 bp
 * 10656 12768: contig of 2113 bp in length
 * 12769 12868: gap of 100 bp
 * 12869 15722: contig of 2854 bp in length
 * 15723 15822: gap of 100 bp
 * 15823 17959: contig of 2137 bp in length
 * 17960 18059: gap of 100 bp
 * 18060 21771: contig of 3712 bp in length
 * 21772 21871: gap of 100 bp
 * 21872 24539: contig of 2668 bp in length
 * 24540 24639: gap of 100 bp
 * 24640 28144: contig of 3505 bp in length
 * 28145 28244: gap of 100 bp
 * 28245 31503: contig of 3259 bp in length
 * 31504 31603: gap of 100 bp
 * 31604 35858: contig of 4255 bp in length
 * 35859 39610: contig of 100 bp
 * 39611 39710: contig of 3652 bp in length
 * 39711 42759: contig of 3049 bp in length
 * 42760 42859: gap of 100 bp
 * 42860 48014: contig of 5155 bp in length
 * 48015 48114: gap of 100 bp
 * 48115 52263: contig of 4149 bp in length
 * 52264 52363: gap of 100 bp
 * 52364 56374: contig of 4011 bp in length
 * 56375 63102: contig of 6628 bp in length
 * 63103 63202: gap of 100 bp
 * 63203 72687: contig of 9485 bp in length
 * 72688 72787: gap of 100 bp
 * 72788 97799: contig of 25012 bp in length
 * 97800 97899: gap of 100 bp
 * 97900 134325: contig of 36426 bp in length
 * 134326 134425: gap of 100 bp
 * 134426 167509: contig of 33084 bp in length.

FEATURES
 source
 1. 167509
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-45K21"

misc_feature
1..1061
/clone_lib="RPCI-11 Human Male BAC"
/note="assembly_fragment"
1162..2381
/note="assembly_fragment"
2482..3771
/note="assembly_fragment"
3872..4949
/note="assembly_fragment"
5050..6119
/note="assembly_fragment"
6220..8023
/note="assembly_fragment"
8124..10555
/note="assembly_fragment"
10656..12768
/note="assembly_fragment"
12869..15722
/note="assembly_fragment"
15823..17959
/note="assembly_fragment"
18060..21771
/note="assembly_fragment"
21872..24539
/note="assembly_fragment"
24640..28144
/note="assembly_fragment"
28245..31503
/note="assembly_fragment"
31604..35858
/note="assembly_fragment"
35959..39610
/note="assembly_fragment"
39711..42759
/note="assembly_fragment"
42860..48014
/note="assembly_fragment"
48115..52263
/note="assembly_fragment"
52364..56374
/note="assembly_fragment"
56475..63102
/note="assembly_fragment"
63203..72687
/note="assembly_fragment"
clone_end:T7
vector_side:right
72788..97799
/note="assembly_fragment"
97900..134325
/note="assembly_fragment"
134426..167509
/note="assembly_fragment"
clone_end:SP6
vector_side:left

ORIGIN
Alignment Scores:
Pred. No.: 2,878-315 Length: 167509
Score: 300.00 Matches: 300
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.32% Indels: 0
DB: 2 Gaps: 0

US-10-066-521-6 (1-1344) x AC012107 (1-167509)

Qy 368 AnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLeuCysVal 387
Db 12664 AACACCGTGAGTCTCTGACACAGTGCAGGTGCCCGCGGTCTCTCATCTGCGTG 12605
Qy 388 AlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThr 407

Db 12604 GCCTCGAGCTGCAGGACGTGGTGGGGAGAGCGTCCGCCCTTCACCAACACCTCACA 12545
Qy 408 GlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgCys 427
Db 12544 GGCCTGCACGCCGCTTTTGTGTTTCATCAGCTCACCTCGAGGGCGTGTCCGGCGTGT 12485
Qy 428 LeuAsnLeuGluArgValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 447
Db 12484 CTCATCTGGAGAAAGAGTTGTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTG 12425
Qy 448 TrpAsnArgGlySerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSer 467
Db 12424 TGGAAATAGGAAGTCAGTGTGTGATGGTGAAGACCTCATGGTTCAAGGACTCGGGAGTCT 12365
Qy 468 GluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyr 487
Db 12364 GAGCTCGTGTCTGTGTTTCATGAAACATCTTCTCCAGACGCCACTGTGAGAGTAC 12305
Qy 488 TyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGlu 507
Db 12304 TACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCCTTGTACTACGTGTAGAG 12245
Qy 508 GlyLeuGluLeuGluProAlaLeuCysProLeuTyrValGluLeuThrLeuLeuSerMet 527
Db 12244 GGCCTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACCTTGAGAAAGACAAAGAGTCCATG 12185
Qy 528 GluLeuLeuGlnAlaGlyPheHisIleHisSerLeuTyrMetLeuLeuLeuPheGly 547
Db 12184 GAGCTTAAACAGGCGGCTTCCATATCCACTCGCTTGGATTGAAGCGTCTCTGTTGGC 12125
Qy 548 LeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeu 567
Db 12124 CTCGTAGCGAAGACGTAAGAGGCCACTGGAGGTCTGTGTGGCTGTCCGCTTCCCTG 12065
Qy 568 GlyValLeuGlnLeuLeuLeuHisTyrValSerLeuLeuGlyGlnGlnProAsnAlaThr 587
Db 12064 GGGGTGAAGCAGAGCTTCTGCACCTGGGTCTCTGTGTGGTTCAGCAGCCCTAATGCCACC 12005
Qy 588 ThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPhe 607
Db 12004 ACCCAGAGACACCTCGAGCGCTTCCACTGTCTTTTCAGAGCTCAAGACAAAGAGTTT 11945
Qy 608 ValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAsp 627
Db 11944 GTTCGCTTGCATTAAACAGCTTCCAAAGAGTGTGGCTCCGATTAAACAGAACCTGCAC 11885
Qy 628 LeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAsp 647
Db 11884 TTGATAGCATCTTCTTCTGCCTCCAGCAGCTCCGTAATTTGCGGAAATTCGGGTGGAT 11825
Qy 648 ValLeuGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTyr 667
Db 11824 GTCAAGGGATCTTCCAGAGATGAGTCCGCTGAGGAGTGTCTGTGGTCCCTCTATGG 11765
RESULT 14
AX427590 1075 bp DNA linear PAT 20-JUN-2002
LOCUS Sequence 3 from Patent WO0232955.
DEFINITION AX427590
ACCESSION AX427590.1 GI:21537712
VERSION AX427590.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nelson,L.M. and Tong,Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 3 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
1..1075
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.87e-203 Length: 1075
Score: 196.00 Matches: 252
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 1
Query Match: 14.58% Indels: 2
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX427590 (1-1075)

```
QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
   |||||
Db 2 CTGGCCTCAGCCCTCGTGCAGAACCGGAGCTTGCACACCTGTGCTCCATCCAAACAGC 61
   |||||
QY 871 LeuGlyAenGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
   |||||
Db 62 CTGGGGAACGAGGTGTAAATCTACTGTGTGCATCCATGAGGCTTCCCACTGTAGTCTG 121
   |||||
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
   |||||
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCTGCACACGGCTGGCTGTGG-TTCTCTTGCACT 180
   |||||
QY 910 uAlaLeuMetGlyAenSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
   |||||
Db 181 TGCCTGTATGGGTAACTCATGGCTGACGCACCTGAGCCTTAGCATGAACCTGTGGAGA 240
   |||||
QY 930 pAsnGlyVallyLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
   |||||
Db 241 CAATGGCTGAAGCTTCTGTGCGAGGTCATGAGAGAACCACTTGTCTATCTCCAGGACCT 300
   |||||
QY 950 uGluLeuVallyCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
   |||||
Db 301 GGAGTTGTGTAATAATGTCATCTCACCGCGCGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
   |||||
QY 970 rArgSerArgHisLeuLeuSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyIva 990
   |||||
Db 361 GAGAGCAGACACCTGAGAGCCCTGGATCTCACGGACAATGCCCTGGGTGACGGTGGGGT 420
   |||||
QY 990 lAlaAlaLeuCysGluGlyLeuLeuGlnLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
   |||||
Db 421 TGCTGCGCTGTGCGAGGACTGAGACAAAGAACAGTGTCTGACGAGACTCGGGTTGAA 480
   |||||
QY 1010 sAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
   |||||
Db 481 GGCATGTGACTGACTTCTGATTGCTGTGAGGCACCTCTCCTTGGCCCTTTCTCTGCAACCG 540
   |||||
QY 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
   |||||
Db 541 GCATCTGACCACTTAAACCTGGTGCAGATAAATCTCAGTCCCAAGGAATGATGAAGCT 600
   |||||
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
   |||||
Db 601 GTGTTGCGCCCTTGCCTGTCCACGCTTAACCTTACAGATAATTGGGCTGTGAAATGCGCA 660
   |||||
QY 1070 nTyrProValGlnIleArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValVa 1090
   |||||
Db 661 GTACCTGTGCAATAAGGAAGCTGTGTGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGT 720
   |||||
QY 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
   |||||
Db 721 AATTGACGGTAGTTGGCATCTCTTTTGATGAGATGACCGG 760
   |||||
```

RESULT 15

AC012107 AC012107 167509 bp DNA linear HTG 28-MAR-2000
DEFINITION Homo sapiens clone RP11-45K21, WORKING DRAFT SEQUENCE, 25 unordered
pieces.
AC012107
AC012107.2 GI:7329252
VERSION

KEYWORDS
SOURCE
ORGANISM

HTG: HTGS PHASE1; HTGS_DRAFT.
Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 167509)

AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens, clone RP11-45K21

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 167509)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArillano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M.,
Ferraira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Gallagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, K., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

TITLE

Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Mar 26, 2000 this sequence version replaced gi:6088020.

All repeats were identified using RepeatMasker:

Smit, A. P. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1180

Center clone name: 45_K_21

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 145437 bases at least Q40

Consensus quality: 159832 bases at least Q30

Insert size: 176000; agarose-fp

Quality coverage: 3.8 in Q20 bases; agarose-fp

Quality coverage: 4.1 in Q20 base.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 25 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1061: contig of 1061 bp in length

* 1062 1161: gap of 100 bp

* 1162 2381: contig of 1220 bp in length

* 2382 2481: gap of 100 bp

* 2482 3771: contig of 1290 bp in length

* 3772 3871: gap of 100 bp

* 3872 4949: contig of 1078 bp in length

* 4950 5049: gap of 100 bp

* 5050 6119: contig of 1070 bp in length

* 6120 8023: contig of 1804 bp in length

* 8024 8123: gap of 100 bp

* 8124 10555: contig of 2432 bp in length

* 10556 10655: gap of 100 bp

* 10656 12768: contig of 2113 bp in length
* 12769 12868: gap of 100 bp
* 12869 15722: contig of 2854 bp in length
* 15723 15822: gap of 100 bp
* 15823 17959: contig of 2137 bp in length
* 17960 18059: gap of 100 bp
* 18060 21771: contig of 3712 bp in length
* 21772 21871: gap of 100 bp
* 21872 24539: contig of 2668 bp in length
* 24540 24639: gap of 100 bp
* 24640 28144: contig of 3505 bp in length
* 28145 28244: gap of 100 bp
* 28245 31503: contig of 3259 bp in length
* 31504 31603: gap of 100 bp
* 31604 35858: contig of 4255 bp in length
* 35859 35958: gap of 100 bp
* 35959 39610: contig of 3652 bp in length
* 39611 39710: gap of 100 bp
* 39711 42759: contig of 3049 bp in length
* 42760 42859: gap of 100 bp
* 42860 48014: contig of 5155 bp in length
* 48015 48114: gap of 100 bp
* 48115 52263: contig of 4149 bp in length
* 52264 52363: gap of 100 bp
* 52364 56374: contig of 4011 bp in length
* 56375 56474: gap of 100 bp
* 56475 63102: contig of 6628 bp in length
* 63103 63202: gap of 100 bp
* 63203 72687: contig of 9485 bp in length
* 72688 72787: gap of 100 bp
* 72788 97799: contig of 25012 bp in length
* 97800 97899: gap of 100 bp
* 97900 134325: contig of 36426 bp in length
* 134326 134425: gap of 100 bp
* 134426 167509: contig of 33084 bp in length.

FEATURES

source
1. .167509
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-45K21"
/clone_lib="RPC1-11 Human Male BAC"
1. .1061
/note="assembly_fragment"
1162. .2381
/note="assembly_fragment"
2482. .3771
/note="assembly_fragment"
3872. .4949
/note="assembly_fragment"
5050. .6119
/note="assembly_fragment"
6220. .8023
/note="assembly_fragment"
8124. .10555
/note="assembly_fragment"
10656. .12768
/note="assembly_fragment"
12869. .15722
/note="assembly_fragment"
15823. .17959
/note="assembly_fragment"
18060. .21771
/note="assembly_fragment"
21872. .24539
/note="assembly_fragment"
24640. .28144
/note="assembly_fragment"
28245. .31503
/note="assembly_fragment"
31604. .35858
/note="assembly_fragment"
35959. .39610

misc_feature
/note="assembly_fragment"
39711. .42759
/note="assembly_fragment"
42860. .48014
/note="assembly_fragment"
48115. .52263
/note="assembly_fragment"
52364. .56374
/note="assembly_fragment"
56475. .63102
/note="assembly_fragment"
63203. .72687
/note="assembly_fragment"
clone_end:T7
vector_side:right
72788. .97799
/note="assembly_fragment"
97900. .134325
/note="assembly_fragment"
134426. .167509
/note="assembly_fragment"
clone_end:SP6
vector_side:left

ORIGIN

Alignment Scores:
Pred. No.: 5.17e-172 Length: 167509
Score: 169.00 Matches: 169
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.57% Indels: 0
DB: 2 Gaps: 0

US-10-066-521-6 (1-1344) x AC012107 (1-167509)

Qy 135 GlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 154
|||
Db 47478 GGACATGGAGGTGACACATGGGACTACAGAGTCAGCTGATGACCAAAATTCGCTGAGGAG 47537

Qy 155 GluAspValArgArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeu 174
|||
Db 47538 GAGATGTAGTCGTGCTAGTTTGAAGAACACATGCTGACTGGCGCGGAAATGCAAAACGTTG 47597

Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
|||
Db 47598 GCTGGTGCITTTTCATTTCAGACCGGTGGGGCTTCGGGCTCGCACGGTGTCTTCGACCGA 47657

Qy 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuGlyTrpAlaGlnGly 214
|||
Db 47658 AAGTCAGGAATGGGAAATCGGCTCTAGCCAGAAAGATCGTGTCTGTCTGGGCGCAAGGT 47717

Qy 215 GlyLeuTyrglnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
|||
Db 47718 GGACTCTACCCAGGGAATGTTCTCCTACGTTCTTCTCCTCCCGTTAGAGAGATGCACGG 47777

Qy 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
|||
Db 47778 AAGAAGGAGAGCAGTGTCTACAGAGTTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCG 47837

Qy 255 ValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAsp 274
|||
Db 47838 GTGACGAGATCATGTCCCGCCAGAAAGGCTGTTGTTTCATCATTCATTCACGCTTCGATGAC 47897

Qy 275 LeuGlySerValLeuAenAspThrLysLeuGlySerAspTrpAlaGluLysGlnPro 294
|||
Db 47898 CTGGGCTCTGTCTCCTCAACAATGACAAAAGCTCTGCAAAAGACTGGGCTGAGNAGCAGCCT 47957

Qy 295 ProPheThrLeuIleArgSerLeuLeu 303
|||
Db 47958 CCGTTTACCCCTCATACGACGTCTGCTG 47984

Search completed: July 16, 2004, 13:35:52

Job time : 12763 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 06:23:54 ; Search time 1107 Seconds
(without alignments)
5157.704 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKSLTSSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745146

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10066521/runat_13072004_122316_11526/app_query.fasta_1.1543
-DB=N Geneseq_29Jan04 -QWTF=fastap -SUFFIX=Oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10066521@cgn 1.1 819 @runat_13072004_122316_11526 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	4035	6 AAL44356	Aal44356 Human PYR
2	1006	74.9	3489	6 AAD41224	Aad41224 Human EMB
3	1006	74.9	3926	7 ADA45218	Ada45218 Human MAT
4	988	73.5	3830	7 ADA45220	Ada45220 Human MAT
5	968	72.0	3226	6 ABX97181	Abx97181 Human NOV
6	968	72.0	5859	6 AAL47135	Aal47135 Pyrin dom
7	968	72.0	6939	6 AAL47131	Aal47131 Pyrin dom
8	968	72.0	6939	6 AAL47140	Aal47140 Pyrin dom

9	616	45.8	3900	6	ABK48628	Abk48628 Human MAT
10	616	45.8	3900	7	AAD49018	Aad49018 Human MAT
11	385	28.6	1157	6	ABK48609	Abk48609 Human MAT
12	385	28.6	1157	7	AAD48999	Aad48999 Human MAT
13	196	14.6	1075	6	ABK48610	Abk48610 Human MAT
14	196	14.6	1075	7	AAD49000	Aad49000 Human MAT
15	12	0.9	3108	6	ADE36416	Ad36416 Human PAA
16	12	0.9	3186	6	AAL44363	Aal44363 Human PYR
17	12	0.9	3300	6	AAL47129	Aal47129 Pyrin dom
18	12	0.9	3306	6	ADC30316	Adc30316 Human nov
19	12	0.9	3447	6	ABK48611	Abk48611 Mouse MAT
20	12	0.9	3447	7	AAD49001	Aad49001 Mouse MAT
21	12	0.9	4931	6	ABL59333	AbL59333 Nucleotid
22	10	0.7	1722	9	ADE62511	Ad62511 Human gen
23	10	0.7	1722	9	ADE62507	Ad62507 Human gen
24	9	0.7	438	5	ABV38053	Abv38053 Human pro
25	9	0.7	578	4	AAL01744	Aal01744 Human rep
26	9	0.7	578	4	ABL97037	AbL97037 Human tes
27	9	0.7	693	6	ABQ40410	Abq40410 Oligonucl
28	9	0.7	693	6	ABQ40524	Abq40524 Oligonucl
29	9	0.7	693	6	ABQ40525	Abq40525 Oligonucl
30	9	0.7	693	6	ABQ40411	Abq40411 Oligonucl
31	9	0.7	702	4	AAK74564	Aak74564 Human imm
32	9	0.7	738	4	AAK59124	Aak59124 Human imm
33	9	0.7	2596	2	AAZ24899	Aaz24899 Human sec
34	9	0.7	4194	5	AAD02761	Aad02761 Human NB-
35	9	0.7	4329	5	AAD02762	Aad02762 Human NB-
36	9	0.7	4422	5	AAD02760	Aad02760 Human NB-
37	9	0.7	4466	5	AAD02765	Aad02765 Human NAC
38	9	0.7	4556	5	AAD02764	Aad02764 Human NAC
39	9	0.7	5100	7	ACC45152	Acc45152 Human cas
40	9	0.7	5122	4	AAD12951	Aad12951 Human G-p
41	9	0.7	5444	7	ACC45143	Acc45143 Human NAC
42	9	0.7	5477	2	AAT33514	Aat33514 Human DEC
43	9	0.7	5523	5	AAF83651	Aaf83651 Human CAR
44	9	0.7	5523	7	AB55497	Ab55497 cDNA enco
45	9	0.7	5523	7	AB556030	Ab556030 cDNA enco

ALIGNMENTS

RESULT 1

AAL44356

ID AAL44356 standard; cDNA; 4035 BP.

XX AAL44356;

AC AAL44356;

DT 31-OCT-2002 (first entry)

DE Human PYRIN-5 cDNA sequence.

XX Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-5.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 1..4035

FT /*tag= a

FT /product= "Human PYRIN-5"

XX WO200261049-A2.

XX 08-AUG-2002.

XX 31-JAN-2002; 2002WO-US002967.

XX 31-JAN-2001; 2001US-0265231P.

PR 10-SEP-2001; 2001US-0318645P.
XX (MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX Bertin J, Wang W, Blatcher M;
XX MPI; 2002-627477/67.
DR P-P8DB; AAO15585.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
PT diagnosing stress-related, apoptotic and inflammatory responses, or for
PT treating inflammatory and immune system disorders, cancers, or
PT neurological diseases.
XX
XX Claim 4; Fig 3; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
CC useful for modulating and diagnosing stress-related, apoptotic and
CC inflammatory responses. The PYRIN protein and DNA sequences are useful
CC for treating: inflammatory disorders and immune system disorders (e.g.
CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
CC (e.g. systemic lupus erythematosus and arthritis); and neurological
CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
CC protein and DNA sequences may also be used in screening assays, detection
CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
CC predictive medicine (e.g. diagnostic assays, clinical trials and
CC pharmacogenomics) and transcription profiling. The present DNA sequence
CC encodes the human PYRIN-5 protein
XX

SQ Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4035
Score: 1344.00 Matches: 1344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAL44356 (1-4035)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAGGAAGAAATTTACAGCAATTCAGGAATTTACTAAAGAAGAAATCTTCAGAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCCACAGTTTGAATCGAAGATGCCAAGTGGAAATGCTTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCTTGCATGATGATTATGAGACATCGTGGCCCTGGGCTACGTCCTCAATTAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysIleSerGln 100
Db 241 AACATGAACCTCGCAACCCCTCTCGGAAGGACACGGGATGACATGAAAAAATTTTCACA 300
Qy 101 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGlnAla 120
Db 301 GCTATGGACAAAGAGTGGCCACAGCAGCAGACAGACAGACAGACAGAAATTTTCAAGCT 360
Qy 121 MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThr 140
Db 361 ATGGAACAAGAAGTGGCCACAGCAGCAGACAGACAGACAGACAGACAGATGGAGGTGACACA 420

Qy 141 TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArgSer 160
Db 421 TGGACTACCAAGAGTCACGTGATGACAAATTCGCTGAGGAGGAGGATGACGTCTGTAGT 480
Qy 161 PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer 180
Db 481 TTTGAAAAACACTGCTGCTGACTGCGCGGAAATGCAAAACGTTGGCTGGTCTTTTATTCA 540
Qy 181 AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLys 200
Db 541 GACCGGTGGGGCTTCGCGGCTCGCACGCGTGGTCTCGCACGAAAGTCAGGAATTTGGAAA 600
Qy 201 SerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLysLeuTyrGlnGlyMet 220
Db 601 TCGGCTCTAGCCAGAAAGGATCGTGTGCTGGCGCAAGGTGGACTCTACCCAGGGAATG 660
Qy 221 PheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSerVal 240
Db 661 TTCTCTACGTCTTCTTCTCTCCCGTTAGAGAGATGACGCGGAAGAGAGACAGTGTCT 720
Qy 241 ThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSer 260
Db 721 ACAGAGTTCACTCCAGGAGTGGCCAGACTCCCGAGGCTCCGCTGACGAGATCATGTCC 780
Qy 261 ArgProGluArgLeuPheIleLeuAspGlyPheAspAspLeuGlySerValLeuAsn 280
Db 781 CGACCAAGAAAGGCTGTGTTCATTCATTCAGCGTTCGATGACCTGGGCTCTGTCTCAAC 840
Qy 281 AsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIleArg 300
Db 841 AATGACACAAAGCTCTGCAAAAGATGGGCTGAGAAGCAGCTCCGTTTCACTCATACGC 900
Qy 301 SerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspVal 320
Db 901 AGTCTGCTGAGGAAGGTCTGCTCTCCCTGAGTCTCTCTGATCGTCCCGTCAGACGTG 960
Qy 321 GlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIle 340
Db 961 GGCACAGAGAAAGCTCAAGTCAGAGGTGCTGTCTCCCGCTTACCTGTGTAGTGAAGAA 1020
Qy 341 SerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThr 360
Db 1021 TCCGGGGAACAAAGAAATCCACTTGTCTTTCAGCGCGGATGCTGAGCATCAGAAGACA 1080
Qy 361 GlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla 380
Db 1081 CAAGGGTGGTGGCATCATGAACAAACCGTGAAGTGTCTGACACAGTCCAGGTGCGCGCC 1140
Qy 381 ValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerValAla 400
Db 1141 GTGGGCTCTCTCATCTGCTGGGCTGACAGTGCAGAGCTGGTGGGGAGAGCGTGGCC 1200
Qy 401 ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro 420
Db 1201 CCCTTCAACCAACCGCTCACAGGCTGACCGCGCTTTTGTGTTCATCAGCTCACCCCT 1260
Qy 421 ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys 440
Db 1261 CGAGGGGTGGTGGCGGCTGTCTCAATCTGGAGAAAGAGTTGTCTCTGAAGCGCTTCTGC 1320
Qy 441 ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet 460
Db 1321 CGTATGGCTGTGAGGAGTGTGGAAATAGGAATCAGTGTGTGATGTGACGACCTCATG 1380
Qy 461 ValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuPro 480
Db 1381 GTTCAAGGACTCGGGAGTCTGAGCTCGTGTCTCTGTTTACATGAACATCTTCTTCCCA 1440
Qy 481 AspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla 500
Db 1441 GACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCC 1500
Qy 501 AlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrVal 520

Db 1501 GCCTTGTAAGTCTAGAGGCGCTGGAATCGAGCCAGCTCTCTGCGCTCTGTACGTT 1560
Qy 521 GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrp 540
Db 1561 GAGAAACAAAGAGGTCATGGAGCTTAAACAGGAGGCTTCCATATCCATCCGCTTTGG 1620
Qy 541 MetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeuGluValLeu 560
Db 1621 ATGAAGCGCTTCTGTTGGCTCGTGAGCGAAGACGTAAGGAGGCCACTGGAGGTCCTG 1680
Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCCGTTCCCTCGGGGTGAAGCAGAGCTTCTGCACTGGGTCTCTCTGTTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTGAGGAGCTTAATGCCACCCAGAGACACCTTGGAGCGCTTCCATCTGCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGCAAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTCCAAAGAGTGTGGCTT 1860
Qy 621 ProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CCGATTAAACCAAGACCTGGACTGTATGATCATCTTCTCTGCTCCAGCACTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAla 660
Db 1921 TTGCGAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 1981 TGTCTGTGGTCCCTCTATGATGGGATTAAGACCTCATTTGAGGAGCAGTGGGAAGAT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 2041 TTCTGCTCCATGTGGCACCAACCACACCTGGCGGAGCTGGACCTGGGCAGCAGCATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 2101 CTGACAGCGCGGCCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCTGCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCTGATGTTTAGAAATGCACAGATTACCCCTGTGTGCAGCACCTCTGGAGATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGGAAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArg 780
Db 2281 GATGTAAGGATGGCGGTGGAAGCCTTAAACACCCCAAAATGTTTGTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGAAATGCTGTGAATTGACCCATGCCCTGTACTGAAGATCTCCCAATCTCTTACGACC 2400
Qy 801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAAGGTGACAGACCAGGAGTAATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTAGTGTGCTTGGAGTCTCCAGAGTCTCCAGTGCCTGCAAGAGCTGATCTGGAGG 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGCACCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCys 880

Db 2581 TTGCACACACCTGTGCTCTATCAACAACAGAGCTGGGGAAACGAAGGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCACCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACACGGCTGGCTGTGGTTTTCTTGCACCTATGGCTTATGGGTAACTCATGGCTCAGCAC 2760
Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluValMet 940
Db 2761 CTGAGCCTTAGCATGAACCCCTGTGGAAGACAATGGCGTGAAGCTTCTGTGCGAGGTCATG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuLysValLysCysHisLeuThrAlaAla 960
Db 2821 AGAAGACCATCTTGTCTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCGCGCG 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGTCTGTGAGAGTCTGTCTGTGTGATCTCCAGGAGCAGACACCTGAAGAGCCTGGATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAAATCCCTGGGTGACGGTGGGGTTGCTGCGCTGTGCGAGGACTGAAGCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGlu 1020
Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGAGCTGACTTCTGTGATGCTGTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCATCTCTTGGCCCTTCTCTGCACCCGCGCATCTGACCACTCTAAACCTGGTGCAGAT 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTCCTGCTCCACGCTCTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuGlu 1080
Db 3181 TTACAGATAATTGGGCTGTGGAATGGCAGTACCTGTGCAAAATGAAGAGCTGCTGGAG 3240
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAAGTCGACTACTCAAGCCCGAGTCGTAATTCAGCGTAGTTGGCATTCTTTTGATGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTCGGCTCCCTGAAAGCCGGCATGGCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysIleArgValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTGTGTGGGGGATGAACCCAGAGCAGAAAGCGTGTGCTGCTTCTGGGTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyLysSer 1160
Db 3421 TTCAAGAGCAGTACACGATTTGCAAGTCTCTCTGCTGGCCACGGCAATTTGTGAGTCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGAGTTGACAAACGTGGAGCAGAGCTCCCGCAACCCCATGGCAGCAGCAACACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTyrSerGlyValaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTGAGTGTGTGATATTCGGAGCCTGGTCTGAAACTGCTGAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyGlyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGCGCTGGATCCAAACAGTGTGATCATGACCAAGAGGTATGGCTGCTCCTAGGAGAGA 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAACAGTGTGATGATGACCAAGCGGTGTGTCTGCTGT 3720

QY 1241 HistTrpGluArgLeuGlySerArgGlyTrpCysLeuAsnSerAlaAspAspHisSerGly 1260
Db 3721 CACTGGGAGCGGCTCTAGGGGCTGGTCTTAACAGTGTGTATGACCACAGCGGT 3780
QY 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAspAsp 1280
Db 3781 GTGTCTTGGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTCTCAACAGTGTGTATGAC 3840
QY 1281 HisSerGlyValAlaAlaTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSer 1300
Db 3841 CACAGCGGTGTGTCTGCTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTGTCAACAGT 3900
QY 1301 AlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlyLeuVal 1320
Db 3901 GCTGATGACCAACAGCGGTGTCTGTGTCACTGGGAGCGGCTGGGCTCGAGGGGCTTGTG 3960
QY 1321 SerAsnSerAlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlu 1340
Db 3961 TCCACAGTGTGTATGACCAACAGCGGTGTGTCTGTGTCACTGGGAGCGGCTGGGCTCGAG 4020
QY 1341 GlyLeuValSer 1344
Db 4021 GGGCTGTGTCT 4032

RESULT 2

AAD41224

ID AAD41224 standard; cDNA; 3489 BP.

XX AC AAD41224;

XX XX

DT 30-OCT-2002 (first entry)

XX DE Human EMBRY-1 cDNA.

XX KW Human; embryogenesis associated protein; AIDS; reproductive disorder;

XX KW infertility; endometriosis; endometrial tumour; inflammatory disorder;

XX KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;

XX KW ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;

XX KW EMBRY-1; allergy; gene therapy; gene; ss.

XX OS Homo sapiens.

XX XX

FH Key

FT CDS

FT Location/Qualifiers

FT 1..3489

FT /*tag= a

FT /product= "EMBRY-1 protein"

XX XX

PN WO200248362-A2.

XX XX

PD 20-JUN-2002.

XX XX

XX 14-NOV-2001; 2001WO-US043956.

XX PF

XX 15-NOV-2000; 2000US-0249407P.

XX PR

XX (INCY-) INCYTE GENOMICS INC.

XX FA

XX Ramkumar J, Arvizu C;

XX PI

XX WPI; 2002-537629/57.

XX DR

XX P-FSDB; AAE25053.

XX DR

XX New polypeptides of human embryogenesis associated proteins for screening

XX FT modulators useful for treating or preventing disorders e.g.

XX PT endometriosis, infertility, allergy, preeclampsia.

XX PT

XX Claim 58; Page 95-96; 97pp; English.

XX FS

XX The invention relates to human embryogenesis associated proteins (EMBRY)

XX CC and nucleic acid molecules encoding such proteins. EMBRY sequences are

XX CC useful for screening modulators useful for treating or preventing

XX CC disorders associated with abnormal expression of EMBRY. The disorders

XX CC

CC treated include reproductive disorders such as infertility,
CC endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,
CC contact dermatitis; disorders of the placenta such as preeclampsia,
CC abruptio placentae etc. Sequences of the invention are also useful for
CC analysing a proteome of a tissue or a cell type. EMBRY proteins are
CC useful as immunogens for preparing antibodies. Polynucleotides of the
CC invention are useful for creating knockin humanised animals or transgenic
CC animals to model human diseases. They are also used in gene therapy. The
CC present sequence is human EMBRY-1 cDNA
XX XX

SQ Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3489
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAD41224 (1-3489)

QY 98 ILeSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGluLe 117

Db 454 ATTTCAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAAGAACAAGAATT 513

QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137

Db 514 TCACAAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAAGAACAAGACATGGA 573

QY 138 GlyAspThrTrpAspTyrIysSerHisValMetThrIysPheAlaGluGluGluAspVal 157

Db 574 GGTGACACATGGGACTACAAAGAGTCAAGTATCACCATAATTCGCTGAGGAGGAGGATGA 633

QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177

Db 634 CGTCGTAGTTTTGAAACACTGCTGCTGAGTCCCGGAAATGCAACAGTTGGCTGGTCT 693

QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyIysSerGly 197

Db 694 TTTGATTACAGCGGTGGGCTTCCGGCTTCGACAGTGGTTCGACGAAAGTCAGGA 753

QY 198 ILeGlyIysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyr 217

Db 754 ATTTGGGAAATTCGGCTCTAGCCAGAAGGATCGTCTGTCTGGGCGCAAGGTGACTCTAC 813

QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgIysLysGlu 237

Db 814 CAGGGAATGTTCTCCCTACGTTCTTCTCCCGTTAGAGAGATGACGGGAAGAAGGAG 873

QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257

Db 874 AGCAGTGTACAGAGTTTCATCTCCAGGGAGTGGCCAGACTCCCGAGCTCCGGTGACGGAG 933

QY 258 IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySer 277

Db 934 ATCATGTCCCGACACAGAAAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993

QY 278 ValLeuAsnAspThrIysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297

Db 994 GTCTCAACAAATGACAAAGCTCTGCAAGAGTGGGCTGAGAAAGAGCTCCGTTACC 1053

QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317

Db 1054 CTCATACGAGTCTGCTGAGGAAGGTCTGCTCCCTGAGTCTCTCTGATGTCACCGTC 1113

QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337

Db 1114 AGAGAGCTGGGACAGAGAAGCTCAAGTCAGAGGTCGTGTCTCCCGTTACCTGTTAGTT 1173

QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357

Db 1174 AGAGGAATCTCCGGGAAACAAGAAATCCACTTGTCTCTTGAGCGCGGATTTGGTGCAGCAT 1233
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 1234 CAGAAGACACAAGGTTGGTGGATCATGAACAACCGTGAGCTGCTCGACAGTGGCCAG 1293
Qy 378 ValProAlaValGlySerLeuLeuLeuCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 1294 GTGCCCGCGTGGGCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAG 1353
Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 1354 AGCGTGGCCCCCTTCAACCAACCGCTCACAGCCCTGCACGCGCTTTTGTGTTTCATCAG 1413
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluAlaArgValValLeuLys 437
Db 1414 CTCACCCCTCGAGGGTGGTGGCGGCTGTCTCAATCTGGAGGAAGAAGTTGTCTCTGAAG 1473
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1474 CGCTTCTGCCGTATGGCTGTGGAGGAGGTGGAATAGGAAGTCAGTGTGTTGACCGGTGAC 1533
Qy 458 AspLeuMetValGlnGlyLeuGlyCysLeuSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 1534 GACCTCATGGTTCAAAGGACTCGGGAGTCTGAGCTCCGCTGCTCTGTTTTCATGAACATC 1593
Qy 478 LeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db 1594 CTTCTCCAGACAGCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTCCAGGAC 1653
Qy 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1654 TTCTGTGCCGCTTGTACTAGTGTGTACAGGGCTGGAAATCGAGGCAGCTCTCTGCCCT 1713
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1714 CTGTACGTTGAAGACAAAGAGGTCCATGGAGCTTAAACAGCGGCGGCTTCCATATCCAC 1773
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 1774 TGGCTTTGGATGAGCGTTTCTTGTGGCTCTGTGAGCGAAGACGTAAGAGGCGCACTG 1833
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 1834 GAGGTCCCTGCTGGGCTGCTCCCTGCGGTGGAAGCAGAGAGCTTCTGCACTGGGTC 1893
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 1894 TCTCTGTTGGTTCAGCAGCCCTTAATGCCACCCACCGAGAGACACCCCTGGACGCTTCCAC 1953
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 1954 TGTCTTTTCGAGACTCAAGACAAGAGTTTGTTCGTTGGCATTTAAACAGCTTCCAAAGAA 2013
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2014 GTGTGGCTTCCGATTAACCAAGAACTGGACITGTAGCATCTTCTTCTGCTCCAGCAC 2073
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2074 TGTCCGTAATTTTCGGGAAATTCGGGTGGATGTCAAAAGGGAATCTTCCCAAGAGATGAGTCC 2133
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluGln 677
Db 2134 GCTGAGGCATGTCTGTGGTCCCTCATGGATGGGGAATAAGACCTCATTTGAGGAGCAG 2193
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2194 TGGGAAGATTTCTGCTCCATGCTTGGCACCCACCCACACCTGCGGAGCTGGACTGGGC 2253
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2254 AGCAGCATCTTGACAGAGCGGGCCAAGAAGACCTGTGTGCCAAGCTGAGGCATCCCAACC 2313

Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2314 TGCAGGATACAGACCCCTGATGTTTAGAAATGACAGATTAATCCCTGGTGTGCAGACCTC 2373
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 2374 TGGAGAATCGTCATGGCCAAACCGTAACCTAAGATCCCTCACTTTGGGAGGACCCACCTG 2433
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2434 AAGGAAGAGGATGTAAAGGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTTGTGGAG 2493
Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2494 TCTTTGAGGCTGGATTGCTGTGGATTGACCCATGCCCTGTTACCTGGAAGATCTCCCAATC 2553
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 2554 CTTACGACCTCCCCAGCCTGAAATCTCTCAGCCTGGCAGGAAACAAGGTGACAGACCG 2613
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2614 GGAATAATGCTCTCAGTGTGCTTGGAGATCTCCAGTGGCCCTGCAGAGCTGATA 2673
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2674 CTGAGAGACTGTGGCATCACAGCCACGGTTCAGAGATCTGGCCTCAGCCCTCGTCAGC 2733
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 2734 AACCGAGCTTGACACACCTGTGCCTATATCCAAACACAGCCTGGGGAACGAAGGTGTAAT 2793
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2794 CTACTGTGTGCATCCATAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2853
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2854 TGCACCTGACAGCGGTGGCTGTGGTTTTCTTGGACTTGGCTTATGGGTAACTCATGG 2913
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 2914 CTGAGCACCTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGC 2973
Qy 938 GluValMetArgGluProSerCysHisLeuGluAspLeuGluLeuValLysCysHisLeu 957
Db 2974 GAGGTCTAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3034 ACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGACAGACCTGGAAGGC 3093
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3094 CTGGATCTCACGACACAATGCCCTGGGTGACGGTGGGTTGCTGCACCTGTCGAGGAGCTG 3153
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3154 AAGCAAAAGAACAGTGTCTGTGACAGACTCGGTTGAAGGCATGTGGACTGACTTCTGAT 3213
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 3214 TGTGTGAGGCATCTCTCTTGGCCCTTCTCTCAACCGGATCTGACACAGCTTAAACCTG 3273
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3274 GTCAGATAAATTCAGTCCCAAGGAATGATGAAGCTGTGTTGGCCTTGGCTTGTGCTCC 3333
Qy 1058 ThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLys 1077
Db 3334 ACCTCTAACTTACAGATAATTTGGGCTGTGGAATGGCAGTACCTGCTGTGCAATAAGGAG 3393

QY 1078 LeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSer 1097
|||||
Db 3394 CTGCTGGAGGAGTGCACCTACTCAAGCCCCAGTGTGTAATTGACGGTAGTGGCATTTCT 3453

QY 1098 PheAspGluAspArg 1103
|||||
Db 3454 TTTGATGAAGATGACCGG 3471

RESULT 3

ID ADA45218 standard; DNA; 3926 BP.
XX
AC ADA45218;
XX

DT 20-NOV-2003 (first entry)

DE Human MATER DNA.

KW human ; MATER; maternal antigen that embryos require; NTP-ase;
KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;
KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;
KW ovarian dysfunction; autoimmune premature ovarian failure;
KW autoimmune disease; contraceptive; autoantibody; ds; gene.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 1. .3489
CDS /*tag= a
FT /product= "MATER"
FT

XX EP1285964-A2.

XX PD 26-FEB-2003.

XX 12-JUL-2002; 2002EP-00090246.

XX 10-AUG-2001; 2001DE-01039874.

XX (SCHD) SCHERING AG.

PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;

XX WPI; 2003-302814/30.

DR P-PSDB; ADA45219.

XX New nucleic acid for human maternal antigen that embryos require protein,
XX useful e.g. in diagnosis and treatment of female infertility.

XX Claim 1; Page 9-11; 31pp; German.

XX This invention describes a novel human MATER (maternal antigen that
XX embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
XX associated with apoptosis, a defect in MATER activity causes growth
XX arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
XX MATER gene is found predominantly in ovary, testis and placenta. The
XX products of the invention have antiinfertility, contraceptive, the
XX antiinflammatory, immunosuppressive and gynaecological activity.
XX Effectors of the MATER polypeptide are used to treat infertility.
XX associated with endometriosis and also ovarian dysfunction, autoimmune
XX premature ovarian failure, inflammation, autoimmune diseases and female
XX infertility, and as contraceptives. Measurements of autoantibodies
XX against MATER in body samples or MATER protein or mRNA in eggs is useful
XX for diagnosis of female infertility, particularly by detecting mutations
XX in the MATER gene with a DNA chip. This sequence encodes the human MATER
XX polypeptide described in the invention.

SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3926
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 7 Gaps: 0

US-10-066-521-6 (1-1344) x ADA45218 (1-3926)

QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGluIle 117
|||||
Db 454 ATTTCAACAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGAGAACAAGAAATT 513
|||||
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137
|||||
Db 514 TCACAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAGAACAAGACATGGA 573
|||||
QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
|||||
Db 574 GGTGACACATGGGACTACAAGAGTCAAGATGACCAATTCGCTGAGGAGGAGGATGTA 633
|||||
QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
|||||
Db 634 CGTCGTAGTTTTTCAAAACACACTGCTGCTGACTGCCCGGAAATGCAAAACGTTGGCTGGTCT 693
|||||
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
|||||
Db 694 TTTGATTACAGCCGCTGGGGCTTCCGGCTCCGACCGTGGTCTGCACGGAAAGTCAGGA 753
|||||
QY 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
|||||
Db 754 ATTTGGAAATCGGCTCTAGCCAGAAGATCGTCTGCTGCTGGCGCAAGGTGACTCTAC 813
|||||
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
|||||
Db 814 CAGGGAATGTTCTCTACGTTCTTCTCTCCCGTTAGAGAGATGCAGCGGAAGAAGGAG 873
|||||
QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
|||||
Db 874 AGCAGTGTACAGATTCTCTCAGGAGTGGCCAGACTCCACAGGCTCCGGTGAACGAG 933
|||||
QY 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
|||||
Db 934 ATCATGTCCCGACCAAGAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
|||||
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
|||||
Db 994 GTCTCAACAATGACACAAGCTCTGCAAGACTGGGCTGAGAAGCAGCCCTCCGTTCACC 1053
|||||
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
|||||
Db 1054 CTCATACGCAGTCTGCTGAGGAAGTCTCTGCTCCCTGAGTCTCTCTGATCGTCAACGTC 1113
|||||
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
|||||
Db 1114 AGACAGCTGGGCACAGAGAAGCTCAAGTCAGAGTGGTGTCTCCCGTTACCTGTTAGTT 1173
|||||
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyHis 357
|||||
Db 1174 AGAGGAATCTCCGGGAAACAAGAAATCCACTTCTCTTGGCGCGGATTTGTTAGCAT 1233
|||||
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
|||||
Db 1234 CAGAAGACACAGAGTTGGCTGGATCATCAACAACCGTGAGTGTCTCGACCACTGCCAG 1293
|||||
QY 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
|||||
Db 1294 GTCCCGCGCTGGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGCTGGTGGGGGAG 1353
|||||
QY 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
|||||
Db 1354 AGCGTCGCCCCCTTCAACCAACCGCTCACAGGCTGCACGCCCTTTTGTGTTTCATCAG 1413
|||||
QY 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
|||||
Db 1414 CTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCCTGAAG 1473

QY 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
DB 1474 CGCTTCGCGGTATGGCTGTGGAGGAGTGTGGAAATAGGAAGTCAGTGTGGCGGTGAC 1533
QY 458 AspLeuMetValGlnGlyLeuGlyGluSerGlnLeuArgAlaLeuPheHisMetAsnIle 477
DB 1534 GACCTCATGGTTCGAAGGACTCGGGAGTCTGAGCTCGGTCTCTGTGTTTCACATGAACATC 1593
QY 478 LeuLeuProAspSerHisCysGluGlyTrpThrPhePheHisLeuSerLeuGlnAsp 497
DB 1594 CTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 1653
QY 498 PheCysAlaAlaLeuTrpThrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
DB 1654 TTCTGTGCGCGCTGTACTAGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCGCCCT 1713
QY 518 LeuTrpValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
DB 1714 CTGTACGTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGCGAGCTTCCATATCCAC 1773
QY 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgAtqProLeu 557
DB 1774 TCGCTTTGGATGAAGCGTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAGAGGCCACTG 1833
QY 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
DB 1834 GAGGTCTGTGCGGTGTGTCCTTCCCTGGGGTGAAGCAGAGCTTCTGCATCGGGTC 1893
QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 1894 TCTCTGTGGTTCAGCAGCCTAATGCCACCAACCCAGGAGACACCTCGGACCGCTTCCAC 1953
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 1954 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGTGTGGCATTAACAGCGCTTCCAAAG 2013
QY 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 2014 GTGTGGCTTCCGATTAACCAAGAACTGGACTGTATAGCATCTTCTTCTGCCTCCAGAC 2073
QY 638 CysProTrpLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB 2074 TGTCCGTATTTGCGAAATTTCCGGTGGATGTCAAAGGATCTTCCCAAGAGATGAGTCC 2133
QY 658 AlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGlnGln 677
DB 2134 GCTGAGGCATGTCTGTGGTCCCTCTATGGATGCGGGATTAAGACCTCATTTGAGGAGCAG 2193
QY 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 2194 TGGGAAGATTTCTGCTCATGCTTGGCACCCACCCACACCTGCGGAGCTGGACCTGGCC 2253
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 2254 AGCAGCATCTCTGACAGACGGCGCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCAAC 2313
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 2314 TGCAGATATACAGACCCCTGTATGTAGAAATGCAAGATTAACCTGTGTGTGTCAGCACCTC 2373
QY 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
DB 2374 TGGAGATCTGTATGGCCACACCGTAACCTAGATCCCTCAACTTGGAGGCGCACCCACTG 2433
QY 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
DB 2434 AAGGAAGAGGATGTAAAGGATGCGGTGTGAAGCCTTAAACACACCCAAATATGTTGTGAG 2493
QY 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIle 797
DB 2494 TCTTTGAGGCTGGATGTGTGGATGTGACCCATCGCTGTACCTGAAGATCTCCCAATC 2553

QY 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
DB 2554 CTTACGACCTCCCCAGCCTGAAATCTCTGAGCCTGCGAGAAACAAGGTGACAGACCAG 2613
QY 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
DB 2614 GGAAGTAATGCTCTCAGTGTGATGCTTTGAGAGTCTCCAGTCCGCTGCAAGAGCTGATA 2673
QY 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
DB 2674 CTGAGAGACTGTGGCATCACAGCCAGGTTGCCAGATCTGGCCTCAGCCCTCGTCAGC 2733
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGAGCTTGACACACCTGTGCTTATCCAAACAACAGCCTGGGGAACGAAGGTGTAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCACCTGACACGCTGCTGTGTTTCTTTCACCTTTCGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTGACGACCTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuIleValLysCysHisLeu 957
DB 2974 GAGGTATGAGAAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
DB 3034 ACCGCGGTGCTGTGAGAGTCTGTCTGTGTATCTCGAGGAGCAGACACCTGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCAGCGACAATGCCCTGGGTGACGGTGGGGTGTCTGCACCTGTGCGAGGACTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCATGTGAGCTGACTTCTGAT 3213
QY 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
DB 3214 TGCTGTGAGCAGCTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTCTAAACCTG 3273
QY 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuLeuCysSerAlaPheAlaCysPro 1057
DB 3274 GTGCAGATTAACCTTCACTCCCAAGGATGATGAGCTGTGTGGCTTGTGCTGTCTCC 3333
QY 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLys 1077
DB 3334 ACCTCTAACTTACAGATAATTTGGGCTGTGGAATGGCAGTACCCCTGTGCAATAAAGAG 3393
QY 1078 LeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSer 1097
DB 3394 CTCTGGAGAAAGTGAGCTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTGTGGCATCT 3453
QY 1098 PheAspGluAspArg 1103
DB 3454 TTTGATGAAGATGACCGG 3471

RESULT 4

ADA45220
ID ADA45220 standard; DNA; 3830 BP.

XX
AC ADA45220;

XX 20-NOV-2003 (first entry)

DE Human MATER splice variant DNA SEQ ID 3.

XX human; MATER: maternal antigen that embryos require; NTP-ase;
 KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;
 KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;
 KW ovarian dysfunction; autoimmune premature ovarian failure;
 KW autoimmune disease; contraceptive; autoantibody; ds; gene.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3432
 FT /*tag= a
 FT /product= "splice variant of MATER"
 XX
 PN EP1285964-A2.
 PD 26-FEB-2003.
 PF 12-JUL-2002; 2002EP-00090246.
 PR 10-AUG-2001; 2001DE-01039874.
 XX (SCHD) SCHERING AG.
 PA
 XX Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
 PI WPI; 2003-302814/30.
 DR P-PSDB; ADA45221.
 XX
 PT New nucleic acid for human maternal antigen that embryos require protein,
 PT useful e.g. in diagnosis and treatment of female infertility.
 XX
 PS Claim 1; Page 17-18; 31pp; German.
 XX
 CC This invention describes a novel human MATER (maternal antigen that
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
 CC associated with apoptosis, a defect in MATER activity causes growth
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
 CC MATER gene is found predominantly in ovary, testis and placenta. The
 CC products of the invention have antiinfertility, contraceptive,
 CC antiinflammatory, immunosuppressive and gynaecological activity.
 CC Effectors of the MATER polypeptide are used to treat infertility.
 CC associated with endometriosis and also ovarian dysfunction, autoimmune
 CC premature ovarian failure, inflammation, autoimmune diseases and female
 CC infertility, and as contraceptives. Measurements of autoantibodies
 CC against MATER in body samples or MATER protein or mRNA in eggs is useful
 CC for diagnosis of female infertility, particularly by detecting mutations
 CC in the MATER gene with a DNA chip. This sequence represents a splice
 CC variant of human MATER in which exon 4 of the usual form is absent.
 XX
 SQ Sequence 3830 BP; 943 A; 946 C; 1041 G; 900 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3830
 Score: 988.00 Matches: 988
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.51% Indels: 0
 DB: 7 Gaps: 0
 US-10-066-521-6 (1-1344) x ADA45220 (1-3830)
 QY 116 GluileSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
 DB 451 GAAATTTTCAAGCATATGGAAACAAGAGGTGCCACAGCAGCAGACAGAGAACAAGGA 510
 QY 136 HisGlyGlyAspThrTpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155
 DB 511 CATGGAGGTGACATATGGGACTACAGAGTCAGTCGTGATGACCAATTCGCTGAGAGGAG 570
 QY 156 AspValArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
 DB 571 GATGTACGTCTAGTATTTTGAACACACTGCTGCTGACTGGCCGGGAAATGCAACGTTGGCT 630

QY 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
 DB 631 GGGTCTTTTATTCACACCGTGGGGCTTCGGGCCCTCGCACGGTGGTTCTGCAAGGAAAG 690
 QY 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
 DB 691 TCAGGAATTTGGGAATCGGCTTAGCCACAGAGATCGTGTCTGGCTGGGCGCAAGGTGA 750
 QY 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
 DB 751 CTCTACAGGGAATGTTCTCTACGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 810
 QY 236 LysGluSerSerValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProVal 255
 DB 811 AAGGAGCAGCAGTGTACAGAGTTCTATCTCAGGGAGTGGCCAGACTCCCAAGGCTCCGGT 870
 QY 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspLeu 275
 DB 871 ACGAGATCATGTCCCGACACAGAAAGCTGTGTTCATCATTTGACCGTTCGATGACCTG 930
 QY 276 GlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPro 295
 DB 931 GGCTCTGTCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAGAAGCAGCCTCCG 990
 QY 296 PheThrIleIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
 DB 991 TTCACCTCATACGAGTCTGTGTGAGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1050
 QY 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
 DB 1051 ACCGTGAGAGCTGGGCACAGAGAGCTCAAGTCAGAGTCTGTCTCTCTCTCTCTCTCT 1110
 QY 336 LeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGly 355
 DB 1111 TTAGTTAGAGGAATCTCCGGGGAACAAGAATCCACTTGTCTCTCTGAGCGCGGATGGT 1170
 QY 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgLeuLeuLeuAspGln 375
 DB 1171 GAGCATCAGAAGACACAGGGTTGCGTGGATCATGAACAACCGTAGCTGCTCGACCCAG 1230
 QY 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
 DB 1231 TGCAGGTGCCCGCTGGGCTCTCTCATCTGGTGGCCCTGCAGCTGCAGGACGTGGTG 1290
 QY 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
 DB 1291 GGGGAGAGCGTCCGCCCTTCAACCAACGCTCACAGGCTGCACGCCGCTTTTGTGTTT 1350
 QY 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValVal 435
 DB 1351 CATCAGTCAACCTCGAGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAGAGTTGTC 1410
 QY 436 LeuLysArgPheCysArgMetAlaValGluGlyValTyrPheAsnArgLysSerValPheAsp 455
 DB 1411 CTGAACGGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGAAATAGGAAGTCAGTGTTCAC 1470
 QY 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
 DB 1471 GGTGACGACCTCATGTTTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTCTGTTTCATG 1530
 QY 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeu 495
 DB 1531 AACATCTCTCCACAGACGCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTC 1590
 QY 496 GlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeu 515
 DB 1591 CAGGACTTCTGTGCCCTTGTACTACGTGTAGAGGGCTTGAATCGAGCCAGCTCTC 1650
 QY 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
 DB 1651 TGCCCTCTGTACGTTGGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCAT 1710

QY 536 IleHisSerLeuTrpMetIysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
DB 1711 ATCCACTCGCTTGGATTGAAGCGTTTCTGTGTGGCTCTCGAGCGAAGACGTAAGGAGG 1770
QY 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
DB 1771 CCACCTGGAGGTCCTGCTGGGTGTCCCGTTCCCTTGGGGGTGAAGCAGAACCTTCTGCAC 1830
QY 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
DB 1831 TGGGTCTCTCTGTGGGTGAGCAGCGCTAATGCCACCCAGCAGACACCCCTGGACGCC 1890
QY 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
DB 1891 TTCACACTCTCTTTTCGAGACTCAAGACAAAGAGTTGTGTGCTTGGCAATTAACAGCTTC 1950
QY 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
DB 1951 CAGAGAGTGGCTTCCGATTAAACAGAACCTGGACTTGATAGCATCTTCTTCTGCGCTC 2010
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
DB 2011 CAGCACTGCTCGTATTTTCGGGAAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGAGAT 2070
QY 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675
DB 2071 GAGTCCGCTGAGGCAATGCTGTGTGCTCTATGGATGCGGATAGACCCCTCATTTGAG 2130
QY 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
DB 2131 GAGCAGTGGGAAGATTTCTGTCTCATGCTTGGCACCACCCAGCCTGCGGCACTGGAC 2190
QY 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
DB 2191 CTGGCAGCAGCATCTGCAGAGCGGGCCATGAAGACCTGTGTGCCAGCTGAGGCAT 2250
QY 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
DB 2251 CCCACCTGCAAGATACAGACCTGATGTGTTAGAAATGCACAGATTACCCCTGTGTGCAG 2310
QY 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThr 755
DB 2311 CACCTCTGGAGAATCGTATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCA 2370
QY 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
DB 2371 CACCTGAAGGAGAGGATGAAGATGGCGTGTGAGCCCTTAAACACCCCAAAATGTTG 2430
QY 776 LeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795
DB 2431 TTGGAGTCTTTGAGCTGGATTGCTGTGATTGACCCATGCCATGCTGTACCTGAAGATCTCC 2490
QY 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThr 815
DB 2491 CAATCTCTTAGACCTCCCCAGCCTGAAATCTCTGAGCCCTGGCAGGAAACAAGTGACA 2550
QY 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
DB 2551 GACCAGGAGATTAATGCCCTCAGTGATGCTTGAAGTCTCCAGTGGCGCTCGCAAG 2610
QY 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
DB 2611 CTGATACTGGAGGACTGTGGCATCACAGCCAGCGGTTGCCAGATCTTGGCTCAGCCCTC 2670
QY 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
DB 2671 GTCAGCAACCGGAGCTTGACACCTGTGCTTATCCAAACACACGCTGGGGAGCAAGGT 2730
QY 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
DB 2731 GTAATCTACTGTGTGATCCATGAGGCTTCCCCACTGTAGCTGTCAGAGGCTGATGCTG 2790
QY 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915

DB 2791 AATCAGTGCACCTGGACACGGCTGGCTGTGGTTTCTTGCACCTTGGCTTATGGGTAAC 2850
QY 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
DB 2851 TCATGGCTGACGACCTTGAGCCTTAGCATGAACCTGTGAAGACAATGGCGTGAAGCTT 2910
QY 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
DB 2911 CTGTGCGAGGTCTATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGT 2970
QY 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValLysSerArgSerArgHisLeu 975
DB 2971 CATCTCACCCCGCGTCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACTG 3030
QY 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995
DB 3031 AAGAGCTGGATCTCAGCGAATGCCCTGGGTGACGGTGGGGTTGCTGCACCTGTGCGAG 3090
QY 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
DB 3091 GGACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACT 3150
QY 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
DB 3151 TCTGATTGCTGTGAGGCACCTCTCTGGCCCTTCTCTGCAACCGGCATCTGACCAGTCTA 3210
QY 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055
DB 3211 AACCTGTGTGAGCAATTAACCTTCAGTCCCAAGGAATGATGAAGCTGTGTCGGCTTTGCC 3270
QY 1056 CysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
DB 3271 TGTCCCACGCTTAACCTTACAGATAATTGGGCTGTGGAAATGGCAGTACCCTGTGCAATA 3330
QY 1076 ArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValLysLeuAspGlySerTrp 1095
DB 3331 AGGAAGCTGTGAGGAAAGTGACAGTACTCAAGCCCGAGTCTGTAATTGACGGTAGTTGG 3390
QY 1096 HisSerPheAspGluAspArg 1103
DB 3391 CATCTTTTGTGAAGATGACCCG 3414
RESULT 5
ABX97181
ID ABX97181 standard; cDNA; 3226 BP.
XX AC ABX97181;
XX DT 20-MAY-2003 (first entry)
XX DE Human NOV125b cDNA.
XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiaesthatic; cancer;
XX KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
XX KW human; gene; ss.
XX OS Homo sapiens.
XX PN WO200272757-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US006908.
XX PR 08-MAR-2001; 2001US-0274101P. *1st clu*
PR 08-MAR-2001; 2001US-0274194P. *no*
PR 08-MAR-2001; 2001US-0274281P. *no*
PR 08-MAR-2001; 2001US-0274322P. *no*
PR 09-MAR-2001; 2001US-0274849P. *no*
PR 12-MAR-2001; 2001US-0275235P. *no*
PR 13-MAR-2001; 2001US-0275578P. *no*
PR 13-MAR-2001; 2001US-0275579P. *no*

13-MAR-2001; 2001US-0275601P. *MD*
 14-MAR-2001; 2001US-0276000P. *MD*
 16-MAR-2001; 2001US-0276776P. *MD*
 19-MAR-2001; 2001US-0276994P. *MD*
 20-MAR-2001; 2001US-0277239P. *MD*
 20-MAR-2001; 2001US-0277321P. *MD*
 20-MAR-2001; 2001US-0277327P. *MD*
 21-MAR-2001; 2001US-0277791P. *MD*
 22-MAR-2001; 2001US-0277833P. *MD*
 23-MAR-2001; 2001US-0278152P. *MD*
 26-MAR-2001; 2001US-0278894P. *MD*
 27-MAR-2001; 2001US-0278999P. *MD*
 27-MAR-2001; 2001US-0279036P. *MD*
 28-MAR-2001; 2001US-0279344P. *MD*
 30-MAR-2001; 2001US-0279738P. *MD*
 30-MAR-2001; 2001US-0279995P. *MD*
 30-MAR-2001; 2001US-0280213P. *MD*
 02-APR-2001; 2001US-0280802P. *MD*
 02-APR-2001; 2001US-0280822P. *MD*
 02-APR-2001; 2001US-0280900P. *MD*
 03-APR-2001; 2001US-0281194P. *MD*
 13-APR-2001; 2001US-0283675P. *MD*
 02-MAY-2001; 2001US-0287424P. *MD*
 02-MAY-2001; 2001US-0288066P. *MD*
 03-MAY-2001; 2001US-0288342P. *MD*
 03-MAY-2001; 2001US-0288528P. *MD*
 15-MAY-2001; 2001US-0291190P. *MD*
 16-MAY-2001; 2001US-0291099P. *MD*
 16-MAY-2001; 2001US-0291240P. *MD*
 30-MAY-2001; 2001US-0294485P. *MD*
 31-MAY-2001; 2001US-0294889P. *MD*
 31-MAY-2001; 2001US-0294899P. *MD*
 18-JUN-2001; 2001US-0299027P. *MD*
 19-JUN-2001; 2001US-0299303P. *MD*
 19-JUN-2001; 2001US-0299310P. *MD*
 10-JUL-2001; 2001US-0304354P. *MD*
 31-JUL-2001; 2001US-0309198P. *MD*
 16-AUG-2001; 2001US-0312903P. *MD*
 10-SEP-2001; 2001US-0318462P. *MD*
 12-SEP-2001; 2001US-0318770P. *MD*
 27-SEP-2001; 2001US-0325430P. *MD*
 27-SEP-2001; 2001US-0325681P. *MD*
 18-OCT-2001; 2001US-0330380P. *MD*
 31-OCT-2001; 2001US-0335301P. *MD*
 14-NOV-2001; 2001US-0332172P. *MD*
 14-NOV-2001; 2001US-0332721P. *MD*
 14-NOV-2001; 2001US-0332722P. *MD*
 14-NOV-2001; 2001US-0333184P. *MD*
 14-NOV-2001; 2001US-0333272P. *MD*
 21-NOV-2001; 2001US-0332094P. *MD*
 03-DEC-2001; 2001US-0337426P. *MD*
 03-DEC-2001; 2001US-0338092P. *MD*
 04-DEC-2001; 2001US-0337185P. *MD*
 03-JAN-2002; 2002US-0345705P. *MD*
 07-MAR-2002; 2002US-00092900. *MD*

(CURA-) CURAGEN CORP.

Padigar M, Spytex KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
 Ferrandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
 Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsbrook JP;
 Lepley DM, Rieger DK;

WPI: 2002-723332/78.
 P-PSDB; ABU65214.

NOVX polypeptides and polynucleotides, useful for preventing or treating
 a disorder associated with aberrant NOVX expression or activity e.g.,
 cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 asthma.

Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
Db TGCCAGGTGCCCGCGTGGCTCTCTCATCTCGGTGGCCCTGCAGCTGCAGAGCTGGTG 1028
Qy 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAAlaPheValPhe 415
Db GGGGAGAGCGTGGCCCTTCAACCAACCGCTCAAGCCCTGCAGCCCTGTCAGCGCTTTTGTGTTT 1088
Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValVal 435
Db CATCAGCTCACCCCTCGAGGGGTGGTCCGGCGCTGTCTCAATCTGGAGAAAGAGTTGTC 1148
Qy 436 LeuIysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
Db CTGAAGCGCTTCTGCGGTATGGCTGTGGAGGAGTGTGGAAATAGGAAGTCAGTGTGTTGAC 1208
Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyCysLeuGluLeuArgAlaLeuPheHisMet 475
Db GGTGACGACCTCATGGTTCAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTTTCACATG 1268
Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTyrPhePheHisLeuSerLeu 495
Db AACATCTCTCTCCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTC 1328
Qy 496 GlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeu 515
Db CAGGACTTCTGCGCCCTTGTAAGTGTGTAGAGGCTGGAATCGAGCCAGCTCTC 1388
Qy 516 CysProLeuTyrValGluIlyThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db TGCCCTCTGTACGTGGAAGACAAAGAGTCCATGGAGCTTAAACAGGCAGGCTTCCAT 1448
Qy 536 IleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
Db ATCCACTCGCTTTGGATGAAGCGTTTCTGTGTTGGCTCGAGCGAAGACGTAAAGAGG 1508
Qy 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
Db CCAGTGGAGGTCTGCTGGCTGTCCCGTTCCTTGGGGTGAAGCAGAGCTTCTGCAC 1568
Qy 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
Db TGGGTCTCTCTGTGGGTTCAGACGCTAAATGCCACCAACCCAGAGACACCTGGAGCC 1628
Qy 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
Db TTCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTC 1688
Qy 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
Db CAAGAAGTGTGGCTTCCGATTAAACGAACTTGACTTGTATGATCATCTTCTTCTGCCTC 1748
Qy 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
Db CAGCACTGTCCGTATTTGCGGAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGAT 1808
Qy 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675
Db GAGTCCCGCTGAGGATGTCTGTGGTCCCTCATATGGATGCGGGATAAGACCTCATTTGAG 1868
Qy 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
Db GAGCAGTGGGAGAGATTTCTGCTCATCTGTGGCACCCACCCACACTTGGCGGAGCTGGAC 1928
Qy 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
Db CTGGGCAGCAGCATCTCTGACAGACGGGCCCATGAAGACCTGTGTGTCGAAGCTGAGGCAT 1988
Qy 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
Db CCCACCTGCAAGATACAGACCTGATGTTAGAAATGCACAGATTACCCCTGGTGTGCAG 2048

Qy 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThr 755
Db CACTCTGGAGATCGTCATGGCCAAACCGTAACCTAAGATCCCTCACTTGGGAGGCACC 2108
Qy 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
Db CACTGAAGGAAGAGGATGTAAAGATGGCGTGTGAAGCTTAAACACACCACAAATGTTTG 2168
Qy 776 LeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuIleSer 795
Db TTGGAGTCTTTGAGGCTGGATGCTGTGGATGTACCCATGCCTGTACCTGGAATCTCC 2228
Qy 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThr 815
Db CAAATCCTTACGACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAAGTGACA 2288
Qy 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
Db GACCAGGAGTAATGCCTCTCAGTGATGCCCTGTAGAGTCTCCAGTGCGCCCTGCAGAAG 2348
Qy 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
Db CTGATCTGGAGACTGTGGCATCACAGCCAGGTTGCCAGAGTCTGGCCTCAGCCCTC 2408
Qy 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
Db GTCAGCAACCGAGCTTGACACACTGTGCTATCCAAACACAGCCTGGGGACGAAGGT 2468
Qy 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
Db GTAAATCTACTGTGTGATCCATGAGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTG 2528
Qy 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915
Db AATCAGTGCCACCTGACACGCTGGCTGTGGTGTTCCTTGCACCTGGCTTATGGGTAAAC 2588
Qy 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
Db TCATGCTGAGCAGCTGAGCCTTAGCATGAAACCTGTGGAGAACATATGGCTGAAGCTT 2648
Qy 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
Db CTGTGGAGGTCTATGAGAGAACCATCTTGTGATCTCCAGGACCTGGAGTTGTTAAAGTGT 2708
Qy 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975
Db CATCTCACCGCCGCTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTG 2768
Qy 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995
Db AAGAGCTGGATCTCACGGACAAATGCTGGGTGACGGTGGGGTGTCTGCACCTGTGGCAG 2828
Qy 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
Db GGACTGAAGCAAAAGAACAGTGTCTTCAGCAGACTCGGGTTGAAGGCATGTGCAGTACT 2888
Qy 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
Db TCTGATGTGTGAGGACCTCTCTTGGCCCTTCTCTCAACCCGATCTCGACCACTGCTA 2948
Qy 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055
Db AACCTGGTCAGANTAACTTCAGTCCCAAGGANTGATGAAGCTGTGTTTGGGCTTTGCC 3008
Qy 1056 CysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
Db TGTCCCACTGCTAACTTACAGATAATTTGGCTGTGGAAATGGCAGTACCTCTGTGCAATA 3068
Qy 1076 ArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrp 1095
Db AGGAAGCTGTGTGAGGAAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTTGACGCTAGTTG 3128
Qy 1096 HisSerPheAspGluAspArgArg 1103

QY	498	PheCysAlaalaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro	517
DB	3589	TTCTGTCCGCTTGTAAGTCTAGAGGGCGCTGGAAATCGAGCCAGGTCTCTGCCT	3648
QY	518	LeuTyrValGluIysThrLysArgSerMetCluLeuLysGlnAlaGlyPheHisIleHis	537
DB	3649	CTGTACGTTGAGAAAGACAAGAGGTCATGAGGCTTAAACAGGCAGGCTTCCATATCCAC	3708
QY	538	SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu	557
DB	3709	TCGCTTTGGATGAAGCGTTCTTGTTGGCTCGTGAGCGAGACGTAAAGAGGCCACTG	3768
QY	558	GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal	577
DB	3769	GAGGTCTGTGGGCTGTCCCGTTCCCTGGGGTGAAGCAGAAGCTTCTGCATCGGCTC	3828
QY	578	SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrIleuAspAlaPheHis	597
DB	3829	TCCTGTGTGGTCAAGCAGCTTAATGCCACACCCAGAGACACCTGGACGCGCTCCAC	3888
QY	598	CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu	617
DB	3889	TGCTTTTCGAGACTCAAGACAAAGAGTTGTTGCTTGGCATTAACACGTTCCAAGAA	3948
QY	618	ValTrpLeuProIleAsnGlnAsnLeuAspLeulleAlaSerSerPheCysLeuGlnHis	637
DB	3949	GTGTGGCTTCGATTAACAGAACCTGGACTTGATAGCATCTTCTCTGCTCCAGCAC	4008
QY	638	CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer	657
DB	4009	TGTCGGTATTTCGGGAAATTCGGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCC	4068
QY	658	AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln	677
DB	4069	GCTGAGGCATGTCTGTGTCCCTCTATGATGCGGATTAAGACCTTCATTGAGAGCAG	4128
QY	678	TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly	697
DB	4129	TGGAGATTTCTGCTCCATGCTTGGCACCCACCCACACCTGGCGGACGTGGACCTGGC	4188
QY	698	SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr	717
DB	4189	AGCAGCATCTGCAGACGCGGCATGAAGACCCCTGTGTGCCAAGCTGAGGCATCCCACC	4248
QY	718	CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu	737
DB	4249	TGCAAGATCAGACCCCTGATGTTTAGAAATGCAAGATTAACCTGCTGTGTGCAGACCTC	4308
QY	738	TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu	757
DB	4309	TGGAGATCTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGAGGACCCACCTG	4368
QY	758	LysGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu	777
DB	4369	AAGGAAGAGATGTAAAGATGGCGTGTGAAGCTTAAACACCCCAAAATGTTTGTGGAG	4428
QY	778	SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrIleuLysIleSerGlnIle	797
DB	4429	TCCTTTAGGCTGGATTCCTGTGGATTGACCATGCCCTGTACCTTGAAGATCTCCAAATC	4488
QY	798	LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln	817
DB	4489	CTTACGACCTCCCCACCTGAAATCTCTGAGCTGSCAGGAAACAGGTGACAGACCAG	4548
QY	818	GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle	837
DB	4549	GGAGTAATGCCTCTCAGTGTATGCCCTTCAGAGTCTCCAGTCTCCAGTCGCGCCTGCA	4608
QY	838	LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer	857
DB	4609	CTGAGGACCTGTGGCATCACAGCCACCGGTTGCCAGAGTCTGGGCTCAGCGCTCGTCA	4668

Qy	858	AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyValAsn	877
Db	4669	AACCGGAGCTTGACACACCTGTGCCTATCCAAACAACAGCCCTGGGGAAACGAAGGTGTAAAT	4728
Qy	878	LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln	897
Db	4729	CTACTGTGTGCATCCATGAGCGTTCCCACTGTAGTCTCGACAGGCTGATGCTGAATCAG	4788
Qy	898	CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTyr	917
Db	4789	TGCCACCTGGACACGCTGGCTGTGTTTCTTTCGCATCTGGCTTATGGGTAACTCATGG	4848
Qy	918	LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyVallysLeuLeuCys	937
Db	4849	CTGACGCACTGAGCCTTAGCATGAACCCCTGTGGAAGACAATGGCGTGAAGCTTCTGTGC	4908
Qy	938	GluValMetArgGluProSerCysHisLeuGlnAspLeuLeuVallysCysHisLeu	957
Db	4909	GAGGTCAATGAGAACCACTTGTGCATCTCCAGGACCTGGAGTTGGTAAAGTGTCACTC	4968
Qy	958	ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer	977
Db	4969	ACGCGCGGTGTGTGAGACTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC	5028
Qy	978	LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu	997
Db	5029	CTGGATCTCACGGACAATGCCCTGGGTGACGGTGGGTTGCTGCACCTGTGCGAGGACTG	5088
Qy	998	LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp	1017
Db	5089	AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGAT	5148
Qy	1018	CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu	1037
Db	5149	TGCTGTGAGCAGCTCTCCTTGGCCCTTCTCTCCAAACGGGCATCTGACCATCTAAACCTG	5208
Qy	1038	ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro	1057
Db	5209	GTGCAGAAATACTTCACTCCCAAGGAATGATGAAGCTGTGTTTCGGCCTTTGCTGTGCC	5268
Qy	1058	ThrSerAsnLeuGlnIlelleGly	1065
Db	5269	ACGCTCAACTTACAGATAATTTGGC	5292
RESULT 7			
AAL47131			
ID	AAL47131 standard; DNA; 6939 BP.		
XX			
AC	AAL47131;		
DT	20-AUG-2002 (first entry)		
XX			
DE	Pyru domain containing protein NALP5/Py8-hs coding sequence.		
XX			
KW	Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;		
KW	antiarteriosclerotic; antiposoriatic; antibacterial; virucide;		
KW	neuroprotective; antiarthritic; antiarheumatic; antiasthmatic;		
KW	nephrotropic; osteopathic; nootropic; intracellular signal transduction;		
KW	inflammation; Alzheimer's disease; infection; psoriasis; asthma;		
KW	arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;		
KW	osteoarthritis; glomerulonephritis; gene; ds.		
OS	Unidentified.		
XX			
FN	W0200240668-A2.		
XX			
PD	23-MAY-2002.		
XX			
PP	30-OCT-2001; 2001WO-EP012545.		
XX			
PR	15-NOV-2000; 2000DE-01056687.		
PR	30-NOV-2000; 2000DE-01059595.		
XX			


```
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuLeuGluGln 677
Db 5149 GCTGAGGCACTGCTGTGCTCCCTCTATGATGCGGATAAGACCCCTCATTTGAGGAGCAG 5208
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 5209 TGGGAAGATTTCTGCTCATGCTTGGCACCCACCCACACCTTGGCGGACGTGACCTGGGC 5268
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 5269 AGCAGCATCTGACNAGCGGGCCATGAAGACCCCTGTGTGCCAAGCTTGAGGCATCCACC 5328
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 5329 TGCAGATACAGACCTGATGTTAGAAATGCACAGATTACCCCTGTTGGTGTCAGCACCTC 5388
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeu 757
Db 5389 TGGAGAATCGTCATGGCCCAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCACCCACTG 5448
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGGAGAGGATGTAAGATGCGGTGTGAACCTTTAAACACCCAAATGTTTGTGGAG 5508
Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGGATTGCTGTGGATTGACCCATGCTGTTACCTGAAGATCTCCCAATC 5568
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTACGACCTCCCGAGCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGCAG 5628
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAATGCTCTCAGTGATGCTTGAAGTCTCCAGTGGCCCTGCGAAGCTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGGAGGACTGTGGCATCACAGCCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGGAGCTTGACACACCTGTGCTTATCCAAACACAGCCTGGGGAACGAAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTGATCATGAGGCTTCCCCACTGTAGTCTGCAAGGCTGATGCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCCACCTGGACACGGCTGGCTGGTTTCTTGCATCTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGACCTTGAGCTTAGCATGAACCTCTGGAAGACATGCGTGAAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 5989 GAGGTGATGAGAAACCAATCTTGTCATCTCAGGACCTGGAGTGGTAAGTGTGTCATCTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCGGTCTGTGAGTCTGTCTCTGTGTATCTCGAGGACGACACCTGGAAGGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 6109 CTGGATCTCAGCGAATATGCCCTGGGTGACCGGTGGGTGTCGACATGTGCGAGGACTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAGAAGACAGTGTCTCAGGACTCGGGTGAAGGCATGTGAGTACTTCTGAT 6228
```

```
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGCACTCTCTTGGCCCTTTCTCTCAACCGGCATCTGACCACTTAAACCTG 6288
Qy 1038 ValClnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGCAGAAATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTTGGCTGTGCC 6348
Qy 1058 ThrSerAsnLeuGlnIleIleGly 1065
Db 6349 ACGTCTAACTTACAGATTAATGGC 6372

RESULT 8
ID AAL47140
AAL47140 standard; DNA; 6939 BP.
XX
AC AAL47140;
XX
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP13/Py17 coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
PN WO200240668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-EP012545.
XX
PR 15-NOV-2000; 2000DE-01056687.
XX
PR 30-NOV-2000; 2000DE-01059595.
XX
PA (APOT-) APOTECH RES & DEV LTD.
XX
XX Tschopp J, Martinon F;
XX
DR WPI; 2002-427093/45.
DR P-PSDB; AAO17868.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
PT proteins and antibodies.
XX
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one pyd (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0
```


US-10-066-521-6 (1-1344) x AAL47140 (1-6939)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 3469 ATTTTCAACAAGCTATGGAACAAGAAGGTGCGACACAGCAGACAGACAGAGAACAAGAAATT 3528

Qy 118 SerGlnAlaMetGluGlnGluGlyAlaAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 3529 TCACAAGCTATGGAACAAGAAGGTGCGACACAGCAGACAGAGAACAAGAAGCATGGA 3588

Qy 138 GlyAspThrTyrAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspVal 157
Db 3589 GGTGACACATGGGACTACAGAGTCAGTGTATGCCAAATTCGCTGAGGAGGAGATGTA 3648

Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTyrProGluMetGlnThrLeuAlaGlyAla 177
Db 3649 CGTCGTAGTTTTGAACAACACTGCTGCTGACTGGCGCGGAATGCAAAAGTTGGCTGGTCT 3708

Qy 178 PheAspSerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 3709 TTTGATTCAGACCGGTGGGCTTCGGCCTCGCACGGTGGTTCTGCAAGGAAAGTCAGGA 3768

Qy 198 IleGlyLysSerAlaLeuAlaAatArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 3769 ATTGGGAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 3828

Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 3829 CAGGGAATGTTCTCTACGCTTCTCTCCCGTTAGAGAGATGCAGCGGAAGAAGAG 3888

Qy 238 SerSerValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGlu 257
Db 3889 AGCAGTGTCAAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCGGTGACGCG 3948

Qy 258 IleMetSerArgProGluArgLeuPheIleAspGlyPheAspAspLeuGlySer 277
Db 3949 ATCATGTCCCGACAGAAAGCTGTGTTTCATCATTCAGCGTTTCGATGACCTGGGCTCT 4008

Qy 278 ValLeuAsnAsnAspThrLysLysCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 4009 GTCCCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAAGCAGCCTCCGTTCA 4068

Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 4069 CTCATACGCAGTCTGCTGAGAAAGTCTGCTCCCTGAGTCTTCTCCATGATCGTCACGCTC 4128

Qy 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 4129 AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGTCTGTGCTCCCGTTTACCTGTTAGTT 4188

Qy 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 4189 AGAGGAATCTCCGGGGAACAAGAATCCACTGTCTCTGAGCGCGGATTTGGTGACAT 4248

Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 4249 CAGAAGACACAAGGTTGCGTGCATCATGTAACAACCGTCAGCTGCTCGACCAAGTGCAC 4308

Qy 378 ValProIaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 4309 GTGCCCGCGTGGGCTCTCTCATCTGGTGGCCCTGCAGCTGCAAGGACGTGGTGGGGAG 4368

Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 4369 AGCGTCGCCCTTCAACCAACGCTCACAGGCTCGACGCGCTTTTGTGTTTCATCAG 4428

Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLys 437
Db 4429 CTCACCCCTCGAGCGTGGTCCGCGCTGTCTCAATCTGAGGAAAGAGTTGCTCTGAAG 4488

Qy 438 ArgPheCysArgMetAlaValGluGlyValTyrAsnArgLysSerValPheAspGlyAsp 457

Db 4489 CGCTTCTCGCGTATGCGTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTCGGTGAC 4548

Qy 458 AspLeuMetValGlnGlyLeuGlyGlySerGluLeuArgAlaAlaLeuPheHisMetAsnIle 477
Db 4549 GACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCGTGTCTGTGTTTACATGAACATC 4608

Qy 478 LeuLeuProAspSerHisCysGluGluTyrTyrPhePheHisLeuSerLeuGlnAsp 497
Db 4609 CTTTCTCCCAAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 4668

Qy 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 4669 TTTCTGTGCCCTTGTACTACGTGTTAGAGGCTCGAATTCGAGCCAGCTCTCTGCCCT 4728

Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 4729 CTGTACGTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGCAGGCTTCCATATCCAC 4788

Qy 538 SerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeu 557
Db 4789 TCGCTTTTGGATGAAGCGTTTCTTTTGGCTCGTGAGCGAAGACGTAAGAGGCGCACATG 4848

Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTyrVal 577
Db 4849 GAGGTCTGCTGGGCTGTCCCGTTCCTGGGGGTGAAGCAGAAAGTCTTGCACTGGGTC 4908

Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 4909 TCTCTGTGGTGCAGCAGCTTAATGCCACACCCAGGAGACACCTTGGACGCTTCCAC 4968

Qy 598 CysLeuPheGluThrGlnAspLysGlyPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 4969 TGTCTTTTCGAGACTCAAGACAAAGATTTGTTGCTTGGCATTAAACAGCTTCCAAGAA 5028

Qy 618 ValTyrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 5029 GTGTGGCTTCGATTAACAGAAACCTGGACTTGATAGCATCTTCTTCTGCTCCAGCAC 5088

Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 5089 TGTCCGTATTTCGGGAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGAGATGAGTCC 5148

Qy 658 AlaGluAlaCysProValValProLeuTyrMetArgAspLysThrLeuIleGluGln 677
Db 5149 GCTGAGGCAATGCTGTGTGCTCTATGATGCGGGATAAGACCTCATTTGAGGAGCAG 5208

Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 5209 TGGGAAGATTTCTGCTCCATGCTTGGACCCACACACCTGGCGCAGCTGGACCTGGGC 5268

Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 5269 AGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACC 5328

Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 5329 TGCAAGATACAGACCTGATGTTTAGAATGTCACGATTACCCCTGGTGTGACGACCTC 5388

Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 5389 TGGAGAAATCGTCATGCGCAACCGTAACCTAAGATCCCTCACTTGGGAGGCACCCACTG 5448

Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGGAAGAGATGTAAGGATGGCGTGTGAAGCCTTAAACACCCAAAATGTTTGTGGAG 5508

Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGATGCTGTGTGATTTGACCATGCTTACCTGAAGATCTCCCAATC 5568

Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTACGACCTCCCGCAGCCTGAAATCTCTGAGCCTGCGAGGAACAAGGTGACAGACCAG 5628

```
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAATGCCTCTCAGTGATGCTTGGAGTCTCCCAAGTCCGCTCCAGAACCTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGGAGGACTGTGGCATCACAGCCAGCGGTTGCCAGAGTCTGGGCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGAGCTTGACACACCTGTGCTATCCAAACAGCCCTGGGAGACGAAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTCATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTCAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCCACCTGGACACGGCTGGCTGTGGTTTCTTCGACTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGACCTGAGCTTAGCATGACCTGTGAGACCAATGGCGTGAAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 5989 GAGGTCATGAGAGAACCATCTTGTCTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTCATCTC 6048
Qy 958 ThrAlaAlaCysGlySerLeuSerCysValLysSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCGCTGCTGAGAGTCTGCTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 6109 CTGATCTCAGCGAATGCTCGGTGACGCTGGGTGGTGTGCTGACCTGTGGAGGAGCTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAAAGAACAGTGTCTCAGCAGACTCGGGTTGAAGGCATGTGACCTCTCTGAT 6228
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGACACTCTCTCTGGCCCTTTCTCTGCAACCGGCACTCTGACCACTGTAACCTG 6288
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGCAGAAATACTTCAGTCCCAAGGAATGATGAGCTGTGTCGGCTTTGCTGTGCTGCTCC 6348
Qy 1058 ThrSerAsnLeuGlnIleIleGly 1065
Db 6349 ACGTCTAACTTACAGATAATTGGC 6372

RESULT 9
ABK48628
ID ABK48628 standard; cDNA; 3900 BP.
XX
AC ABK48628;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER cDNA.
XX
KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POP;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3603
/*tag= a
```

```
FT XX /product= "Human MATER"
PN XX WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
DR WPI; 2002-454595/48.
DR P-PSDB; AAU79526.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 11; Page 80-85; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA.
CC Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to
CC published human genomic DNA sequences, while the remainder of the
CC sequence was determined by direct cloning of human ovarian cDNAs
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: Gaps: 0

US-10-066-521-6 (1-1344) x ABK48628 (1-3900)
Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 568 ATTTCAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAACAGAAATT 627
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 628 TCACAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGACAGACATGGA 687
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACACATGGGACTACAAAGACTCAGTATGACCAAAATTCGCTGAGGAGGAGTGA 747
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 748 CGTCGTAGTATTTGAAACACACTGCTGCTGACGCGGAAATGCAACGTTGGCTGGTGTCT 807
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
```

|||||
808 TTTGATTTCAGACCGGTGGGGCTTCGGGCTCCGACGGTGTCTTCGACGGAAAGTCAGGA 867
Qy |||||
198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db |||||
868 ATTGGGAAATCGGCTTAGCCAGGAAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 927
Qy |||||
218 GlnGlyMetPheSerTyrValPhePheLeuP-roValArgGluMetGlnArgLysLysGlu 237
Db |||||
928 CAGGGAATGTTCTCTACGCTTCTCTCCCGGTTAGAGAGATGCGAGCGGAAGAAGGAG 987
Qy |||||
238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db |||||
988 AGCAGTGTACAGAGATTCTCCAGGGAGTGGCCAGACTCCCAAGGCTCCGGTACCAGG 1047
Qy |||||
258 IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db |||||
1048 ATCATGTCCCGACACAGAAAGGCTGTGTTTCATCATTCGACGGTTCGATGACCTGGGCTCT 1107
Qy |||||
278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPropheThr 297
Db |||||
1108 GTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAGAGACGACCTCCGTTCAAC 1167
Qy |||||
298 LeuIleArgSerLeuLeuAtgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db |||||
1168 CTCATACGAGTCTGCTGAGGAAGGTCTGTCTCCCTGAGTCCCTTCCTGATCGTCAACGCTC 1227
Qy |||||
318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db |||||
1228 AGACAGCTGGGCACAGAGAGCTCAAGTCAAGAGGTGTGTCTCCCGCTTACCTGTTAGTT 1287
Qy |||||
338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db |||||
1288 AGAGGAATCTCCGGGGAACAAAGAAATCCACTGTCTCTTGAGGCGGGATGGTGAGCAT 1347
Qy |||||
358 GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysG1 377
Db |||||
1348 CAGAAAGACACAGGGTGGCTGCATCAT-CAACACCGTGAGCTGTGCACACGTCGCCA 1406
Qy |||||
377 nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyG1 397
Db |||||
1407 GGTGCCCGCCGTGGCTCTCTCATCTCGTGTGCCCTGCAGCTGCAGGACGTGGTGGGGA 1466
Qy |||||
397 uSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db |||||
1467 GAGCGTCGCCCCCTTCAACCAACACGCTCACAGGCTGCACGCCCTTTTGC-GTTTCATC 1525
Qy |||||
417 lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db |||||
1526 AGCTCACCCCTCGAGGGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAGAGTGTCTCTGA 1585
Qy |||||
437 ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA 457
Db |||||
1586 AGCGCTTCTGCGGTATGGCTGTGAGGAGGTGGGAATAGGAAGTCAGTGTGTTGATGGTG 1645
Qy |||||
457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db |||||
1646 ACGACCTCATGTTTCAAGGACTCGGGAGTCTGAGCTCCGTGTCTCTTTTCACATGAACA 1705
Qy |||||
477 leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA 497
Db |||||
1706 TCCTTCTCCACAGACGACCTGTGAGGAGTACTACACTTCTTCCACTCAGTCTCAGG 1765
Qy |||||
497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db |||||
1766 ACTTCTGCGCGCTTGTACTACTGTGTAGAGGCTTGGAAATCGAGCCAGCTCTCTGCC 1825
Qy |||||
517 roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db |||||
1826 CTCGTGACGTGTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGAGGCTTCATATCC 1885
Qy |||||
537 isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL 557
Db |||||

Db |||||
1886 ACTCGCTTTGGATGAAGCGTTCCTTGTGGCTCGTGAGCGAAGAGTGAAGAGAGCCAC 1945
Qy |||||
557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV 577
Db |||||
1946 TGGAGGTCTCTGTGGCTGTCCCGTTCCTCGGGGTGAAGCAGAGAAGCTTCTGCACCTGG 2005
Qy |||||
577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db |||||
2006 TCTCTCTGTGGGTACGACGCTTAATGCCACCAACCCAGGAGACACCTGGAGCCCTTCC 2065
Qy |||||
597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db |||||
2066 ACTGTCTTTTTCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGATTCAAG 2125
Qy |||||
617 luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db |||||
2126 AAGTGTGGCTTCGATTAACACAGAACCTGACTTGTATAGCATCTTCTTCTGCTCCAGC 2185
Qy |||||
637 isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS 657
Db |||||
2186 ACTGTCCGTATTTGCGGAAATTCGGGTGGATGTCAAGGGGATCTTTCCCAAGAGATGAGT 2245
Qy |||||
657 erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG 677
Db |||||
2246 CCCTGAGGCACTGTCTGTGTCCCTCTATGATGCGGGATAGACCCCTCATTTGAGGAGC 2305
Qy |||||
677 lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db |||||
2306 AGTGGGAAGATTCTCTCATGTCTGGCACCCACCAACCTGTGCGGACGTGGACCTGG 2365
Qy |||||
697 lysSerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT 717
Db |||||
2366 GCAGCAGCATCTTCAGACGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGACATCCCA 2425
Qy |||||
717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db |||||
2426 CTTGCAAGATACAGACCTGTATGTTTGAATATGCACAGATTACCCCTGGTGTGCACACC 2485
Qy |||||
737 euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db |||||
2486 TCTGGAGAATCGTCAATGCGCAACCGTAACTAAGATCCCTCAACTTGGGAGGCACCCACC 2545
Qy |||||
757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db |||||
2546 TGAAGGAAGAGGATGAAGATGGCGTGTGAAGCCCTTAAACACACCCAAATTTTGTGG 2605
Qy |||||
777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db |||||
2606 AGTCTTTGAGGCTGGATTGCTGTGGATTGCCCATGCTGTACCTGAAGATCTCCCAA 2665
Qy |||||
797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db |||||
2666 TCCTTACGACCTCCCGCAGCTCAAAATCTCTGAGCCTGGCAGGAAACAAGGTGCACAGAC 2725
Qy |||||
817 lnClyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuLysLeu 836
Db |||||
2726 AGGAGGTAAAC-GCCTCTCAGTGATGCTTGGAGGTCTCCAGTGCGCCCTTGCAGAAAGCTG 2784
Qy |||||
837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db |||||
2785 ATACTGGAGGACTGTGGCATCACAGCCAGGGTTGCCAGAGTCTGGCCTCAGCCCTCGTC 2844
Qy |||||
857 SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal 876
Db |||||
2845 AGCAACCGGAGCTTGACACACCTGTGCCTATCCAAACAACAGCCTGGGGAAACGAAGGTGA 2904
Qy |||||
877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db |||||
2905 AATCTACTGTGTGCATCCATGAGGCTTCCCACTAGTGTCTGCAGAGGCTGTATGCTGAAT 2964
Qy |||||
897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db |||||
2965 CAGTGCACCTGGACACGGCTGGCTGTGG-TTCTCTTGCACTTGGCTTATGGCTTAACATC 3023

```
QY 916 rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyVallysLeuLe 936
Db 3024 ATGGCTGACGACCTTAGCATCAACCTGTGAGACAAATGCGGTGAAGCTTCT 3083
QY 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuVallysCysHi 956
Db 3084 GTGCAGGTCTATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGTC 3143
QY 956 sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy 976
Db 3144 TCTACCCGCGGTGCTGTGAGAGTCTCTCTGTGTATCTGAGGACACACCTGAA 3203
QY 976 sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluG1 996
Db 3204 GAGCCTGGATCTCACGGACAATGCCCTGGGTGAGGTGGGTGGTGGCTGGCGAGGG 3263
QY 996 yLeuLysGlnIlysAnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe 1016
Db 3264 ACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTC 3323
QY 1016 rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs 1036
Db 3324 TGATTGCTGTGAGGACATCTCTTGGCCCTTCTTGCACCGGCATCTGACCACTCTAAA 3383
QY 1036 nLeuValGlnAnAnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCy 1056
Db 3384 CCTGGTGCAAGATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGCTG 3443
QY 1056 sProThrSerAnLeuGlnIleGlyLeuTrpLysTrpGlnTrpProValGlnIleAr 1076
Db 3444 TCCCACTGTGAGGACATCTCTTGGCCCTTCTTGCACCGGCATCTGACCACTCTAAA 3503
QY 1076 gLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHi 1096
Db 3504 GAAGCTGTGGAGGAGTGCAGCTACTCAGCCCGAGTGTGAATGACGTAATGACGTAATG 3563
QY 1096 sSerPheAspGluAspArg 1103
Db 3564 TTCCTTTCATGAAGATGACCGG 3585
RESULT 10
AAD49018
ID AAD49018 standard; cDNA; 3900 BP.
XX AC AAD49018;
XX DT 07-MAR-2003 (first entry)
XX DE Human MATER cDNA.
XX KW Human; MATER protein; infertility; fertility; contraceptive agent;
XX KW gene therapy; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..3603
XX FT /*tag= a
XX FT /product= "Human MATER protein"
XX PN WO200281492-A1.
XX PD 17-OCT-2002.
XX PF 29-MAR-2002; 2002WO-US009776.
XX PR 04-APR-2001; 2001WO-US010981.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Nelson LM, Tong Z;
XX XX
```

```
DR WPI; 2003-058494/05.
DR P-PSDB; AAE31749.
XX
PT New isolated variant MATER proteins and nucleic acids, useful for
PT diagnosing, prognosing and treating infertility and reduced fertility,
PT and as contraceptive agents.
XX
PS Example 2; Page 93-98; 110pp; English.
XX
CC The present invention relates to novel MATER proteins and polynucleotides
CC encoding such proteins. The MATER proteins are essential to fertility.
CC Sequences of the invention are useful for diagnosing, prognosing and
CC treating infertility, reduced fertility and as contraceptive agents. They
CC are also useful in gene therapy. The method is useful for detecting a
CC predisposition to or pre-symptomatic screening of an individual for
CC infertility or reduced fertility. The present sequence is human MATER
CC cDNA
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: 7 Gaps: 0

US-10-066-521-6 (1-1344) x AAD49018 (1-3900)
QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 568 ATTTCAACAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGAGAAGAAGAAATT 627
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 628 TCAACAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGAGAAGAAGACATGGA 687
QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACACATGGGACTACAAAGAGTCAAGAGTCAAGTATGACCAAAATTCGCTGAGGAGGAGTGA 747
QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 748 CGTCTGATGTTTGAACAACACTGCTGCTGACTGCGCGGAATGCAACGTTGGCTGCT 807
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 808 TTTGATTACAGCCGTGGGGCTTCCGGCTTCGCACGCTGCTTCGCACGGAAGTCAGGA 867
QY 198 IleGlyLysSerAlaLeuAlaArgAlaValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 868 ATTTGGGAATTCGGCTCTAGCCAGAAGATCGTCTGTCTGGCGGCGAAGTGGACTCTAC 927
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 928 CAGGGAATGTTCTCTACGCTCTTCTCTCCCTCCCTAGAGAGATGACGCGAAGAGGAG 987
QY 238 SerSerValThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 988 AGCAGTGTACAGAGTTCATCTCCAGGAGTGGCCAGACTCCCGAGGCTCCCGTGCAGGAG 1047
QY 258 IleMetSerArgProGluArgLeuLeuPheIleAspGlyPheAspLeuGlySer 277
Db 1048 ATCATGTCCCGACCAAGAGGCTGTTGTTTCATCATTTGACGGTTCGATGACCTGGGCTCT 1107
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlnLysGlnProProPheThr 297
Db 1108 GTCTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGGACGCTCTGCTTACC 1167
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1168 CTCATACGCACTGCTGCTGAGGAAGGTCTTCTCTCCCTGAGTCTTCTCTGATCGTCCGCTC 1227
```

QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1228 AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTGCTGTCTCCCGCTTACCTGTTAGTT 1287
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHis 357
Db 1288 AGAGGAATCTCCGGGAAACAAGAATTCACCTTGCTTGGCCGGGATGGTGACAT 1347
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysG1 377
Db 1348 CAGAAGACACAGAGGTTGCGTGCATCAT-CAACAACCGTGAGCTGTCGACAGTGCCA 1406
QY 377 nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyG1 397
Db 1407 GGTGCCCGCGTGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGA 1466
QY 397 userValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db 1467 GAGCGTGCCTTCAACCAACGCTCACAGGCTGCAGCCGCTTTGCG-GTTTCATC 1525
QY 417 lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db 1526 AGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGGAGGAAGAGTTGTCTCGA 1585
QY 437 ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA 457
Db 1586 AGCCCTTCTGCGGTATGGCTGTGAGGAGGTGGGAATAGGAATGCTAGTGTTCATGGTG 1645
QY 457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db 1646 ACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCAATGAA 1705
QY 477 leLeuLeuProAspSerHisCysGluGluTrpThrPhePheHisLeuSerLeuGlnA 497
Db 1706 TCCTTCTCCACAGACGACCTGTGAGGAGTACTACACCTTCTCCACCTCAGTCTCCAGG 1765
QY 497 spPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db 1766 ACTTCTGCGCGCTTGTACTACGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCC 1825
QY 517 roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db 1826 CTCTGTACGTTGAGAACAAAGAGTCCATGGAGCTTAAACAGCAGGCTTCCATATCC 1885
QY 537 isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL 557
Db 1886 ACTCGCTTTGGATGAAGCGTTTCTGTTTGGCTCGTGAGCGAAGACGTAAGGAGGCCAC 1945
QY 557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV 577
Db 1946 TGGAGGTCTGCTGGGTGTCCTGCTCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGG 2005
QY 577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db 2006 TCTCTGTGGGTTCAGACGCTAATGCCACCCACCCAGAGACACCTCGACGCTTCC 2065
QY 597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db 2066 ACTGCTTTTCGAGACTCAAGACAAGAGTTTGTTCGCTTGGCATTTAAACAGACTTCCA 2125
QY 617 luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db 2126 AAGTGTGGCTTCCGATTAACAGAACTGGACTGTAGATCATCTTCTTCTGCTCCAGC 2185
QY 637 isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS 657
Db 2186 ACTGTCGTATTTCCGGAATTTCCGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGT 2245
QY 657 erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG 677
Db 2246 CCGCTGAGGATGTCTGTGTCCTCTATGATGCGGGATAGACCCCTCATTTGAGGAGC 2305

QY 677 lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db 2306 AGTGGGAAGATTTCCTCATGCTTTGGCACCCACCACACCTTGGCGCAGCTGGACTGG 2365
QY 697 lYSerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProI 717
Db 2366 GCAGCAGCATCTCGACAGAGCGGCCCATGAAGACCCTGTGTGCCAAGCTGAGGCATCCCA 2425
QY 717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db 2426 CCTGCAAGATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAACACC 2485
QY 737 euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db 2486 TCTGGAGAATCGTCATGGGCAACCCGTAACCTAAGATCCCTCAACTTGGGAGGCACCCACC 2545
QY 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db 2546 TGAAGGAAAGGATGTAAAGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTTGTGG 2605
QY 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db 2606 AGTCTTTGAGGCTGGATTGCTGTGATTGACCATGCTGTGTACCTGCTGAAGATCTCCCAA 2665
QY 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCCTTACGACCTCCCCAGCTGAAATCTGTAGCCTGCGAGAAACAAGGTGACAGACC 2725
QY 817 lnGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGAGGTAAC-GCCTCTCAGTGATGCTTTGAGGGTCTCCAGTGCGCCCTGCAGAGCTG 2784
QY 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db 2785 ATACTGGAGACTGTGGCATCACAGCACGGGTGCGAGAGTCTGGCCTCAGCCCTCGTC 2844
QY 857 SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal 876
Db 2845 AGCAACCGAGCTTGACACACCTTGCTCCATCCACACAGCCTGGGGAACGAAGGTGTA 2904
QY 877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db 2905 AATCTACTGTGTCGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAAT 2964
QY 897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db 2965 CAGTGCACCTGACACGCGCTGCTGTGG-TTCTCTTGCACTTGGCTTATGGGTAACTC 3023
QY 916 rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe 936
Db 3024 ATGGCTGACGCACTGAGCCTTAGCATGAACCTGTGGAAAGACAATGGCGTGAGCTTCT 3083
QY 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi 956
Db 3084 GTGCGAGGTCATGAGAGAACCATTCTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCA 3143
QY 956 sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy 976
Db 3144 TCTCACCGCGGTGCTGTGAGAGTCTGCTGTGTGATCTCGAGGAGCAGACACCTGAA 3203
QY 976 sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluG1 996
Db 3204 GAGCCTGGATCTCACGGACAAATGCCCTGGGTGACGTGGGTGGGTGCTGCCCTGTGCGAGGG 3263
QY 996 yLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe 1016
Db 3264 ACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTC 3323
QY 1016 rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs 1036
Db 3324 TGATTCGTGTGAGGCACTCTCCTTGGCCCTTTCTCCAAACCCGCACTCTCACCAGTCTAA 3383
QY 1036 nLeuValGlnAsnAsnPheSerProLysGlyMetLysLeuCysSerAlaPheAlaCy 1056

Db 3384 CTGTGTGAGATAACTTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCTTTGGCTGT 3443
Qy 1056 sProThrSerAsnLeuGlnIleilegLyLeuTrpLysTrpGlnTyrProValGlnIleAr 1076
Db 3444 TCCACGCTTAACCTACAGATTAATTTGGGCTGTGGAAATGGCAGTACCTCTGTGCAATTAAG 3503
Qy 1076 gLysLeuLeuGluValGlnLeuLysProArgValValIleAspGlySerTrpHi 1096
Db 3504 GAAGCTGCTGGAGGAAGTGCAGCTACTCAAGCCCGAGTGTGTAATTTGACGGTAGTTGCA 3563
Qy 1096 sSerPheAspGluAspAspArg 1103
Db 3564 TTCTTTTGATGAAGATGACCCG 3585

RESULT 11

ABK48609
ID ABK48609 standard; cDNA; 1157 BP.

XX AC ABK48609;

XX DT 13-AUG-2002 (first entry)

XX DE Human MATER cDNA fragment #1.

XX KW Human; gene; ss; contraceptive; antiinfertility; MATER;

KW Maternal antigen that embryos require; WATER null phenotype; oocyte;

KW early embryonic survival; premature ovarian failure; POF;

XX autoimmune infertility; chromosome 19; gene therapy; fertility.

XX OS Homo sapiens.

XX FT Key Location/Qualifiers
XX CDS 1..1155

XX FT /*tag= a

XX FT /product= "Human MATER"

XX FT /partial

XX FT /note= "No start or stop codon shown"

XX WO200232955-A1.

XX PD 25-APR-2002.

XX PF 04-APR-2001; 2001WO-US010981.

XX PR 18-OCT-2000; 2000US-0241510P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Nelson LM, Tong Z;

XX DR WPI: 2002-454595/48.

XX DR P-PSDB; AAU79523.

PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.

PS Claim 11; Page 68-69; 93pp; English.

CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy Maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein

CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA
CC fragment #1, which was isolated from a human cDNA library

XX SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 1157
Score: 385.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.65% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x ABK48609 (1-1157)

Qy 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleAspGly 271

Db 1 CAAGCTCCGGTGACGGAGATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGACGGT 60

Qy 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291

Db 61 TTCGATGACCTGGGCTCTGTCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAG 120

Qy 292 LysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311

Db 121 AAGCAGCTTCGGTTACCCCTCATACGAGTCTGTCTGAGGAAGGCTCTGCTCCCTGAGTCC 180

Qy 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331

Db 181 TTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGCTGTGTCT 240

Qy 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGlu 351

Db 241 CCCCGTTACCTGTAGTTAGAGGAATCTCCGGGGAACAAAGAAATCCACTTCTCTCTTGG 300

Qy 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371

Db 301 CGCGGGAATTGTGAGCATCAGACAAAGGTTGTGCGCATCATGAACAACCGTGAG 360

Qy 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391

Db 361 CTGCTCCAGCAGTCCAGGTGCCCGCTGGGCTCTCTCATCTCGGTGGCCCTCAGCGTG 420

Qy 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411

Db 421 CAGGACGTGGTGGGGAGAGCGTCGCCCTTCAACCAACCGCTCACAGGCTGCACGCC 480

Qy 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431

Db 481 GCTTTTGTGTTCATCAGCTCACCCCTCGAGGCGGTGTTCGGGCGCTGTCTCAATCTGGAG 540

Qy 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451

Db 541 GAAAGAGTTGTCCTGAAGCGCTTCGCGGTATGGCTGTGGAGGGAGGTGTGGAAATAGGAG 600

Qy 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471

Db 601 TCAGTGTTCAGCGGTGACGACCTCATGGTTCAGGACTCGGGGAGTCTGAGCTCCGTGCT 660

Qy 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePhe 491

Db 661 CTGTTTCAATGAAATCTCTCTCCAGACAGCCACTGTGTGAGGAGTACTACACCTTCTTC 720

Qy 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIle 511

Db 721 CACCTCAGTCTCCAGGACTCTGTGCGGCTCTGTACTACGTGTAGAGGGCTGGAATC 780

Qy 512 GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531

Db 781 GAGCCAGCTCTCTGCCCTCTGTACGTTGAGAGAACAAAGAGGTCCATGGAGCTTTAAACAG 840


```
QY 512 GluProAlaLeuCyProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCCAGCTCTCTGCCCTCTGTACGTTGAGAGACAAAGAGGTCATGGAGCTTAAACAG 840
QY 532 AlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTGTGGCTCTGTGAGCGAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GAGCTAAGGAGCCACTGGAGGTCCTGTGGGCTCTCCGTTCCCTGGGGGTGAAGCAG 960
QY 572 LysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGCATGGGTCTCTGTGTGGTCAAGCAGCCTAATGCCACCACCCAGGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGGAGCGCTTCCACTGCTTTTCGAGACTCAAGACAAGAGTTTGTTCGCTTGGCA 1080
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1081 TTAACAGCTTCCAGAGAGTGTGGCTTCGATTAAACAGACCTGGACTTGATGCACT 1140
QY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCCTCCAG 1155

RESULT 13
ABK48610
ID ABK48610 standard; cDNA; 1075 BP.
AC ABK48610;
DT 13-AUG-2002 (first entry)
DE Human MATER cDNA fragment #2.
KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 2..778
FT /tag= a
FT /product= "Human MATER"
FT /partial
FT /note= "No start codon shown"
XX
FN WO200232955-A1.
XX
XX 25-APR-2002.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX 18-OCT-2000; 2000US-0241510P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2002-454595/48.
DR P-PSDB; AAU79524.
XX
XX New isolated human Maternal Antigen That Embryos Require protein and
XX polynucleotide, useful in diagnosing or treating fertility or reduced
XX fertility, or as a contraceptive.
PS Claim 11; Page 70-71; 93pp; English.
XX
```

```
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency, a
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA
CC fragment #2, which was isolated from a human cDNA library
XX
SQ Sequence 1075 BP; 278 A; 251 C; 281 G; 265 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,46e-184 Length: 1075
Score: 196.00 Matches: 252
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 1
Query Match: 14.58% Indels: 2
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x ABK48610 (1-1075)
QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
Db 2 CTGGCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGCTTATCCACACAGC 61
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAAGGTGTAATCTACTGTGTCATGAGGCTTCCCACTGTAGTCTG 121
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCTGGACACGGCTGGCTGTGG-TTCTCTTGGCACT 180
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TGCCTTATGGGTAACTCATGGCTGACGCACCTGAGGCTTAGCATGAACCTGTGGAGA 240
QY 930 pAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCAAGAGAGAACCATCTTGTTCATCTCCAGGACCT 300
QY 950 uGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTTGGTAAATGTCTATCCACCGCGCGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
QY 970 rArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGAAGAGCTGGATCTCAGCGACAAATGCCCTGGGTGAGCGTGGGT 420
QY 990 lAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
Db 421 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGAAAGACAGTGTCTGACGAGACTCGGGTTGA 480
QY 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCATGTGGACTGACTTCTGATTCTGTGAGGCACTCTCTCTTGGCCCTTTCTCTCAACCG 540
QY 1030 qHisLeuThrSerLeuAsnLeuValGlnAsnAspPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACAGTCTAAACCTGGTGCAAGATAACTTCACTCCCAAGGAATGATGAAGCT 600
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
```

Db 601 GTGTTGGCCCTTGCTGCTCCACGCTCTAACTCAGATAAATGGGCTGTGGAATGGCA 660
Qy 1070 nTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArgValVa 1090
Db 661 GTACCTGTGCAATAAGGAAGCTGCTGGAGGAAGTGACGCTACTCAAGCCCGAGTGT 720
Qy 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTTGGCAATCTTTTGATGAAGATGACCG 760

RESULT 14

AAAD49000
ID AAD49000 standard; cDNA; 1075 BP.

XX AAD49000;

XX 07-MAR-2003 (first entry)

XX Human MATER cDNA fragment 2.

XX Human; MATER protein; infertility; fertility; contraceptive agent;
KW gene therapy; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..778

FT /product= "Human MATER protein fragment 2"
FT /transl_except= (pos:200..202, aa:Thr)
FT /transl_except= (pos:650..652, aa:Thr)
FT /transl_except= (pos:656..658, aa:Thr)
FT /transl_except= (pos:734..736, aa:Thr)
FT /transl_except= (pos:764..766, aa:Thr)
FT /transl_except= (pos:767..769, aa:Thr)
FT /note= "No start codon"
FT /partial

XX W0200281492-AL.

XX 17-OCT-2002.

XX 29-MAR-2002; 2002WO-US009776.

XX 04-APR-2001; 2001WO-US010981.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Nelson LM, Tong Z;

XX WPI; 2003-058494/05.

DR P-PSDB; AAE31747.

XX New isolated variant MATER proteins and nucleic acids, useful for
PT diagnosing, prognosing and treating infertility and reduced fertility,
PT and as contraceptive agents.

PS Example 2; Page 81-82; 110pp; English.

XX The present invention relates to novel MATER proteins and polynucleotides
CC encoding such proteins. The MATER proteins are essential to fertility.
CC Sequences of the invention are useful for diagnosing, prognosing and
CC treating infertility, reduced fertility and as contraceptive agents. They
CC are also useful in gene therapy. The method is useful for detecting a
CC predisposition to or pre-symptomatic screening of an individual for
CC infertility or reduced fertility. The present sequence is human MATER
CC cDNA fragment 2

XX Sequence 1075 BP; 278 A; 251 C; 281 G; 265 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.46e-184 Length: 1075

Score: 196.00 Matches: 252

Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 1
Query Match: 14.58% Indels: 2
DB: 7 Gaps: 0

US-10-066-521-6 (1-1344) x AAD49000 (1-1075)

Qy 851 LeuAlaSerAlaIleuValSerAsnArgSerIleuThrHisLeuCysLeuSerAsnAsnSer 870
Db 2 CTGGCCCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGCTTATCAACAACAGC 61
Qy 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAAACGAAGGTGTAATCTACTGTGTCGATCCATGAGGCTTCCCCACTGTAGTCTG 121
Qy 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCCACCTGGACACGGCTGGCTGGTGG-TTCTCTTGCACT 180
Qy 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TGGCGCTTATGGGTAACCTCATGGCTGACGCACCTGAGCCTTAGCATGAACCTCTGGAAGA 240
Qy 930 pAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGGAGGTCATGAGAAACCACTCTGTCTATCTCAGGACCT 300
Qy 950 uGluLeuValLysCysHisLeuThrAlaLaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAATTGGTAAATGTCATCTCACCGCCGGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
Qy 970 ArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGAAGACCTTGGATCTCACGGACAATGCCCTGGGTGACGGTGGGGT 420
Qy 990 lAlaAlaLeuCysGluGlyLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
Db 421 TGCTGGCTGTGGAGGACTGAAGCAAAAGACAGTGTCTGACGAGACTCGGGTTGAA 480
Qy 1010 sAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCATGTGGACTGACTTCTGATTGTGTGAGGCACCTCTCTTGGCCCTTTCTTCTGCAACCG 540
Qy 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetIleLysLe 1050
Db 541 GCATCTGACCAGTCTAAACCTGGTGAGAGATAAATCTTCAGTCCCAAGGAATGATGAAGCT 600
Qy 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTpGl 1070
Db 601 GTGTTGGCCCTTTGCCCTGCCACGTCCTAACTTACAGATAAATTGGGCTGTGGAAATGGCA 660
Qy 1070 nTyrProValGlnIleArgLysLeuGluValGlnLeuLysProArgValVa 1090
Db 661 GTACCTGTGCAAAATAAGGAAGCTGTCTGGAGGAAGTGACGCTACTCAAGCCCGAGTCTGT 720
Qy 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTTGGCAATCTTTTGATGAAGATGACCGG 760

RESULT 15

ADE36416

ID ADE36416 standard; cDNA; 3108 BP.

XX ADE36416;

XX AC

XX 29-JAN-2004 (first entry)

XX Human PAAD and nucleotide binding protein PAN6 cDNA.

XX cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
KW antiallergic; antiulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;

